

FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCGTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCCGGCCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC
CGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCCTTC
CTGGAGGAGCAGCTGGGGTCTTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGG
CTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCCCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCCTCA
GCTGAGGGAAGGTACGAGTTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAAATAAAAATGAAA
CGTGAAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTGCGCCCTCGCCTCCCCGAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCACGGCGGGGACTATGGTGAAATTCCCGGCGCTCACGCACTACTGGCCCCGTGATC
CGGTTCTTGGTGCCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGCCTGGTGTTT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCTAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTCACGACATCATCCC
TGACAGAAGTGCCCGGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCCGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCCCATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGCCCCAGCTCTGTGCTGCG
GATCATCGTCTCATCGCCAGCCTCGTGGTCCTACCTACCTGGGGGTGCACGGTGCGACCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACCTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTGTTTTGTTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCCTCACGGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAAACTCGGCTTCCTTTGATTTGCTTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCCGGAGTCCCCGTTACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACCTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCACTTCTACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTCACAAAA
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCTGAAGGCTTGCTTTTCCCTCGC
CTTTCTGAAGGTGCGATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT
TACAGTGAACCTGAAGCTTTAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCACACCATAACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACTATGTGGTCCTCCCTGAAATAACGCATTTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTGGAAAGCAGGTTTTTTCTTTTAAAAAATTATAGACACGTTCACT
AAATTGATTTAGTCAGAATTCCTAGACTGAAAGAACCTAAACAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCTATTTTCG
ATTTTCAATAAAATGTCTCTAATACAAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSDFKNVGLVVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDES
GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCLGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFELCVVPLR
IFSFFPVPVTVRAHLTGWMLTLKKTFFVLAPSSVLRIIVLIASLVVLPYLGVHGATLGVGSL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPPTEEVTDIVEMRENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

059250-1337

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

09978259-103504

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTTTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTCCCTTTGCGGATTTTCT
CCTTTTTCCAGTTCCAGTCACAGNGAGGGCGCATCTACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTTCCCTCCTGGCGGGCA

0978259-104501

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTTACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGCGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCTGCGGCTCCTTCTCTTCA
TCCTCATCCAGCTGGTGTCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCATTTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTCACTACTACACTGAGCCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCCTGCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCTAT
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTTGCCAACCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCTCCTCCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACCTCCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGT
CCTGGTCACGTCCCCCAGGGGACCTGCCCCCTTCTGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAACAAGCCAGTGCGTGTAACAAAAA

FIGURE 9

MGACLGACSLLLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SOLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSW
NQRWLKGAEEDCSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTF
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVA
ACEGRAFDNEQDGVITYSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICAS
WAGLLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCGAGTCTTTGCTGCCGAAGCTG
TGA CTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTATCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGA ACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAAATTGCAGGA
GCTGTCACGGGAAGTCTTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACCTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAAACAAGACAAGGACTTGAAAGTGCTCTGAACTTGAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTTCTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAIEYHNRFDVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

09578290 "03501
"05"0" 66282660

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTCGTTTCATG
GCTGGCGCCGAACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

FIGURE 14

GAGCCGCCGCCGCGCGCGCGCGCGCACTGCAGCCCCAGGCCCGGCCCGCCCCACCCACGTCT
GCGTTGCTGCCCCGCTGGGCCAGGCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTGCGGGG
GCTTCGCGTGTTCCAAGAAGTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCAATTGCAGTGGGCATCTTCTTGTTCTTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTCAAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAACTGCT
GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCCGCGCAATCCTAGTGCATTCTTTTGAATGAGAAAACAAGGAAGAT
TTCCTTTCGTATTATGATCTTGTTTCACTTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTTCTTCCGTTGCTGAAAAATATTTGAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCAATTGTGGGCACTGTCCACTGTGGCCTT
TCTTAGCATTTTTTACCTGCAGAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCTAC
TGAAAAAGAGTGGAATTTATTAAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCCAATTTTTTTTTGGTCTTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTGTGGTTTAAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGCTAA
AAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTCTGTAAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAAAGAAAGTAATGGAAG

[illegible]

><subunit 1 of 1, 204 aa, 1 stop

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRVFGGIGLFFSFTEILGVWL
TYRYRNQKDPANPSAFL

amino acids 1-34

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCTGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

0973239-101501

FIGURE 17

AATCCCAAATTCCCCAATTTTTTTGGNCTTTTATAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCCATAAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCTTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

09078295 " 201501

FIGURE 19

CAGTCACCAATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTGTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTTCAAGTGAACCTTCCACCTGATTGTGTCCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCTTGGGATCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTGAGACAAAGTTGCCCCCTGCAGAGGTGAGCTGCCCGC
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTTCAGAAGATCACTCCGGGTGCTACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACCTGCTCCTGA
GGAGGCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCCTCCTCGGTCACTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAAGTAAACAGTTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTGTCCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCAGTGTTTTTGTAGATAATGTAGTTAGGTGAGTGTAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTTGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCAATTATACTTGGGGGGTTGGGGGATGTTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCACAAACGAAAAATAATCTGGCCCAAATGTGAGTTGTACTGAGTTTGAGAAA
CCCCAGCCTAATGAAACCCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTGAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCACAAAAACAGCTGTCGCCAAACACCG
ACTCTGTCGTTGCCTTGATCTTGAACCTTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGLVWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRCQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSAAPPTLNPAPOKSAAPGTAPPEEAPGGLPPPPPTPSSDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

099222 01501

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTTGTGAAGTTTGTG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATAACCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTTTGACCCGT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACCTTCCCAGAATCTGGGCAACAACCTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCCTTGCCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCCCTGGATAGCCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGCTTCTGGTACTCCTCTCTAAATAACAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLLGHLLTVDDTYGRPILEVPESVTGPWKGDVNLPCITYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRHLVSHKVPGDVSLQLSTLEMDDRSYHTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRLSLQCQARGSPPISYIWKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTTMT
YPLKATSTVKQSWDWTDDMDGYLGETSAGPGKSLPVFAIILIIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

05973300-10501

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCQESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

09978290-101503
"06284660"

[illegible]

AAAAAAAAAAAA

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594

><subunit 1 of 1, 655 aa, 1 stop

><MW: 71845, pI: 8.22, NX(S/T): 8

MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIQTYRHVDRATG
QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGTLPFSSTSPSPGTAIFPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTQNREKWIYYCNGHGIDILKLVAQAQVGSQWKDIY
QFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPSPSPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLL
RCDSTSSGSSALSRRNGSFITKEKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

05973290-101501

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTGGCC
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAAGATGCCGTGCTCCAGGTGTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCACTGCACAGCCT
TGGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCACAGGGCTACCACCTGTGCGGGGGCTCTGTATCAC
GCCCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGTTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGA
ACTTCCCCGATGGAAGTGTGCTGGACGTGAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCCTGTCTGAACACGCGGCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTCACCTCCTTCCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCT**GA**
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCTTGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCGAAAGAGGCACCCCTTCCATCTGATTCCAGCACAACCTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGAGGTGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCCAACTAATTTTTGTATTTTTAGTAGAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGTGTTGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTCCCCTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGAAAATTCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCCAGAAGTGCAGAACTGCAGTC
ACTGCACGTTTTTCATCTCTAGGGACCAGAACCACCAACCCACCTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTTCTT
TGTAGCATTTGGTGTCTGACGTATTATTGTCTTTGATTCCAAATAATATGTTTCCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKV TALHH
SVYVREGCASGHVVT LQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIIITAAHCVYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYH SKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHA AVPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSFGIGCAEVNKP GVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 29

CCCACGCGTCCGTCTTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAAAGGCTGAGCTTGT
 GCCAGAACGGCGCGCGCGCGCGCACGACGACACACGCGGGGAACTTTTTTAAAAATGAAAGGCTAGAAGA
 GCTCAGCGCGCGCGCGCGCGCGTGCAGGAGGCTCCGAGAGTGAAGTGCAGCGCTCGCCGCGCGGCGCGAGAGCTGCTGCACTGAAG
 CGCCCGGCGCGCGCTCGCGCGCGCGTGGGATGGTGCAGCGCTCGCCGCGCGGCGCGAGAGCTGCTGCACTGAAG
 GCGGCGAGCAGATGGCAGCGCGCGCGCTGCGCGTGTCCCCGCGCGCGCGCTCCCTGCTCGCCCTGGCCGGTGTCT
 GCTCGCGCGCTGCGAGGCGCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTGAGTGCCTCTGT
 TCGGAGTGGGGACCTCTGGATCCCACTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGTGAATATTTCGACT
 ACAACGGGAAAGCAAAGAAGTATCATAAATCTGGAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAC
 CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
 ACGGGGATATTCTGATTCAGCAGTCACTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTTGTGTTTGAAAATGA
 AAGCTATGTCTTAGAACCAATGAAAAGTGAACCAACAGATACAACTCTTCCAGCGAAGAAGCTGAAAAGCGT
 CCGGGGATCATGTGGATCACATCACACACCAACCTCGCTGCAAGAATGTGTTTCCACCACCTCTCAGAC
 ATGGGCAAGAAGGCATAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
 AGAGTTTTCAGAGGCAAGGAAAAGATCTGGAAGAAAGTAAAGCAGCGATTAATAGAGATTGCTAATCACGTTGACAA
 GTTTTACAGACCACTGAACATTCCGATCGTGTGGTAGGCGTGAAGTGTGAATGACATGGACAAATGCTCTGT
 AAGTCAGGACCCATTACCAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
 TGACAATGCGCAGCTTGTGAGTGGGGTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG
 CACGGCAGACCACTCTGGGGGAATTTGTATGGACCATTCAGACAATCCCTTGGTGCAGCGGTGACCTGGCACA
 TGAGCTGGGCGACAATTTCCGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCCGGTTGAGAA
 AGGAGGCTGCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTT
 GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCTGTTTAACTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCA
 GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
 CTGCAATGCCACCCTGTACCTGAAGCCGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
 GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC
 TCACTGCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTGAGGATGTGGACGGCTACTGCTACAATGGCAT
 CTGCCAGACTCACGAGCAGCAGTGTGTGACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA
 GAGAGTCAATTTCTGAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCTTTTGCCAAATGCGAGAT
 GAGAGATGCTAAATGTGGAAGAAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTC
 CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGATTCTGTGCGGGGGACCCAGTGTACTTGGGCGATGA
 CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGAGATGGAAAAATCTGCCTGAATCGTCAATGTCA
 AAATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCACGGCAGAGGGGTGTGCAACAACAGGAAGAA
 CTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCC
 CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG
 ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCAGTG
 CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCCCTCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
 TCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCAGCTCTGCAGAAA
 CAGTGAAGGAAGGGCAGCGACTTCTGGTTGAGCTTCTGCTAAACATGGACATGCTTCAGTGTCTGCTCCTGAG
 AGAGTAGCAGGTTACCCTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTC
 ACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCAGTGTCCCCCTTCCCCAGTGACACCTCAGCCT
 TGGCAGCCCTGATGACTGGTCTCTGGGTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTATTATATGAAAAT
 AGCAGGGTTTTAGTTTTTAAATTTATCAGAGACCTGCCACCCATTCCATCTCCATCCAAGCAAACCTGAATGGCAA
 TGAAACAACTGGAGAAGAGGTAGGAGAAAGGGCGGTGAACCTCTGGCTTTTGTGTGGACATGCGTGACCAGC
 AGTACTCAGGTTTGGGGTTTGAGAAAGCCAGGGAACCCACAGAGTCACCAACCCCTTCATTTAAACAAGTAAGAA
 TGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT
 GAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSSASVRSGDLWIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGDLEKVKQRLIEIANHVDFK
YRPLNIRIVLVGVEVWMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRCCNATTCTLKPDCAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICFERVNSAGDPYGN
CGKVS KS SFAKCEMRDAKCGKIQCGGASRPVIGTNAVSIETNIPLOQGGRI LCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRQCONISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPIRQAEARQEAAESNRERGGQGPVGSQEHASTASLTLLI

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAACGAAA
ACAAAACAGTTTTTGGGGGTTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTTCCTGGTTGAGNTTTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT
GGGCCCAGTGTCCCCTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCGCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTTAAGAAGTTAAT
GAAACCATAACCTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTTCGAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCATACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCCGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVLLAL
LHLYH

Signal peptide:

amino acids 15-27

06973249-121201

FIGURE 34

GCCGCGGCGAGAGCGCGCCAGCCCCGCGCGATGCCCCGCGCGCCAGGACGCCTCCTCCCGTGCTGGCCCCGGC
 CGGCGGCCCTGACTGCGTGCTGCTGCTGCTGCTGGGCCATGGCGGCGGCGGGCGCTGGGGCGCCCCGGGCCAGG
 AGGCGGCGGCGGCGGCGGCGGACGGGCCCGCGGCGAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCACC
 TGTACACGGCCGACATGTTACGCACGGGATCCAGAGCGCCGCGCACTTCGTGATGTTCTTCGCGCCCTGGTGTG
 GACACTGCCAGCGGCTGCAGCCGACTTGGGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT
 ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCCAGGGGGTGCAGAGGATACCCACCTTAA
 AGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCTCGGGACTTCAGACACTGGAAAACTGGATGC
 TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGAACCGCCAGTGCCCCGAGCTCAAGCAAG
 GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCCTCGCTC
 CGTGGTGTGGTCACTGCAAAGCCCTGGTCCAACCTGGGAGCAGCTGGCTCTGGGCCTTGAACATTCCGAAACTG
 TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCCGGAAACCAGGTTCTGTGGCTATCCCACTC
 TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGGAAAGCGGGATTGGAGTCACTGAGGGAGTACG
 TGGAGTCGCAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCGGGTGTGGCAG
 CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGCAGAAGGAA
 TAACCTTCATCAAGTTTTATGCTCCATGGTGTGGTCATTGTAAGACTCTGGCTCCTACTTGGGAGGAACTCTCTA
 AAAAGGAATTCCCTGGTCTGGCGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAACGGAATATCTGCAGCA
 AGTATTCGGTACGAGGCTACCCACGTTATTGCTTTTCCGAGGAGGGAAGAAAGTCACTGAGCACAGTGGAGGCA
 GAGACCTTGACTCGTTACACCGCTTTGTCTGAGCCAAGCGAAAGACGAACCTTAGGAACACAGTTGGAGGTAC
 CTCTCCTGCCCAGCTCCCGCACCCCTGCGTTTAGGAGTTCACTCCCACAGAGGCCACTGGGTTCCAGTGGTGGCT
 GTTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTCCAAAGCCAACACACTCTACAG
 ATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACCTCATGGTCACTGTGTAAACATTTTCAGTGGCGATA
 TATCCCCCTTTGACCTTCTCTTGATGAAATTTACATGGTTTTCTTTGAGACTAAAAATAGCGTTGAGGGAAATGAAA
 TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC
 CCACGAGTTCTGGAAAGGTGGCCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTCAAGTTGACTCAATAC
 TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC
 TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
 GTGCCTGAGTAAGAGAATGCTGATGCCATAACTTTATGTGTGATACCTTGTCAAATCAGTTACTGTTGAGGGGAT
 CCTTCTGTTTCTCACGGGGTGAAAATGCTTTAGTTCTCATGTTAACACGAAGCCAGAGCCACATGAACTGT
 TGGATGTCTTCTTAGAAAGGGTAGGCATGGAAAATTCACGAGGCTCATTCTCAGTATCTCATTAACTCATTGA
 AAGATTCAGTTGTATTGTACCTGCGGTGACAAGACCAGACAGGCTTTCCAGGCCTGGGTATCCAGGGAGGC
 TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCCTTTAGAGG
 CTTGCTATACCTTGGTCTGCTTCAAGGAGGTGACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC
 CAAAGACAGATGTCAGTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTA
 ACTCATGCTGTCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA
 GATAGGTGTTTTGTCTTTTACCATCGAGCTACTTCCATAATAACCACTTTGCATCCAACACTCTTCACCACCT
 CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAAGTGGTGGTAGGAATCTTAGAAAACAAGACCATTATA
 CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCACG
 TATGGTTACAGATAATTCTTTTTTAAAAAAACCAACCTCCTAGAGAAGCAACTGTCAAGAGTCTTGTACA
 CAAACTTCAGCTTTGCATCACGAGTCTTGTAATTCAGAAATCAAAGTGGTACAATTTGTTTGTACACTAT
 GATACTTTCTAAATAAATCTTTTTTTTTTAA

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLTYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELSASNFEHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESLEHYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDLHFRFVLSQAKDEL

Signal sequence:

amino acids 1-32

09978259 " 7071501
199707 6528660

FIGURE 36

CTTTTCTGAGGAACCACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCGGGTGTTTATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTTCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPEGEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGCRY
RKFBVGQLDISIARLKTSMKFVKNIAGIRETEEFYYIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

05978299-101301

FIGURE 38

GGTTCTATCGATTCTGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCGTGTGGAGTGGTGGGGCCTGGGTGGGAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACTGGCCCAGCAGCGCGCCGCGC
ACACCTTTCTCATTCACGGCTCGCGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG
CGGCGAGGGGAGCGCTGAGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGGCCGAGGGGACGGTGCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTCACCTGGAGCAACTGTGGCGCTGCTCCTCCCGCTGGCCCAGAGTTTCTGTGGCTCTG
GTTCCGGGCTGGCCAAGGCCGGCCTGCGCACTGCCTTTGTGCCACCCGCCCTGCGCCGGGCC
CCCTGCTGCACTGCCTCCGCGAGCTGCGGCGCGCGCGCTGGTGCTGGCGCGCAGAGTTTCTG
GAGTCCCTGGAGCCGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCTTCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC
TCTACCACATGTCCGGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCACTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTGGGGCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCTGGCTTTACAAGC
ATATCTTCCCCTTCTCCTTGATTGCTATGATGTCAACACAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCCAGAGCTGGCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCC
GCCCCGATTCCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCCTACCTGCCCCCTCACAACTGCCCGGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTGAGAAGTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGG
CCGTTGCAGGTGTACTGGGCTGTCAGGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCACTTGTTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQQRTRAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFVRALCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRRFSSYSEAERESNRAARAFRLALGWDWGPDPGGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGLAKAGLRRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKIILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTGPIRDPQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPPELAQ GK
LLKDVFVRPGDVFFNTGDLVLCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATAACACTCGCTCTC
GCTTGTCCATCTCCCTCCGGGGGAGCCGCGCGCGCTCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACCTCGGATTGCAGCTCTGAACCC
CCATGGTGTTTTTTTAAACACTTCTTTTTCTTCTCTTCTCCTCGTTTTTGATTGCACCGTTTTCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG
TTGCACCATGCCTTCTTGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCCGCGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTTCAGCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTTGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTTAGCTGATAAACCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTTCAGGTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGTTCAGCCCAACCCAGGGTGATCCG
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTGCGGGGCTTCCCACTGTGAGGCCCTGCA
ACAATACTGTCTCAACGTATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTCAGGGATGTGGTTCAGCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCACGGAGTTTGAGTTTGTACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTGGCACTGCAGAGACTGTGCAGATA
ATCTTGGGTTTTTGGTCAGATGAAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACTTACCGTTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACCGGGTGCCAGACTGAACTG
CTTCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTTCGTACCAGGAGATTTTCTTACCTTCATTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTTCTCATTTAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMF
VRTYGMLYMQNSEVFQDLFTELKRYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEY
LECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSQKVFQCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHASKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

0978290 "701203
701203 0978290

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCCTCACTTCGCCGCTGGTGAGTGTGCGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGAGAAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA
AAAGTTCCTGTCCACTGTGATTCTCAATTCCTTGCTTGGTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTTCTTTTTTTTTTTTCTTGGTGGAAGCTGCTCTAGGGAGGGGGGAGGAGGAGGAGAAAGTGAAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTCTTCCGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTACCCTCTGTGTGGGTGCATG
TGTGCGCCCCGACGCGCGCGGGGCGCGTGGTTCTCCGCGTGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA
GGGGCTGTGCGGGGCATCCGCCTCCGCCTTCTCCACAGGCCTGTGTCTGTCTGCTGGAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTTGGGGCCAGGCCTTAGAAGAGGAGG
AAGAAGGGGCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCA
TTTTCATCTAGCGGATGATCAGGGATTAGAGATGTGGGTACCACGGATCTGAGATTAAACACCTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAGTATCAGATACACACCGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTTTAC
CTCTGGACAATGCCACCCTACCTCAGAACTGAAGGAGGTTGGATATTCAACGCATATGGTTCGGAAAATGGCACT
TGGGTTTTTAACAGAAAAGAATGCATGCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACTACAAATGTGACAGTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCAT
ACCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCACTCACCACTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAAGTGGCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTTCACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGCATAACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTCAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCTACGGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCGTCAGGTATCCCCCAAAGACCCAGAAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAAGCAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAAGCAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTTATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCCACTGAAAACTGTCTGTCTCAGTG
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCTGAGTTCCACTGCTGTGCTTCAGTCAACTGACCAAACACTGCTTTGAATTATAGGAGGAGAAACAATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAAACTACCTTTGATAAATTAC

[illegible][illegible]

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296

><subunit 1 of 1, 515 aa, 1 stop

><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHGSEIKTPTLDKLAAGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMGVKGWHLGFNRKEC
MPTRRGFDFTFFGSLGSGDYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLA
LKTYGFYNNSSIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

U97899-40501
"0000" 000000

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTCGGGAACGCGGCCAGTGCAGGCATCACGGGTTGTTAGCATCGGCACGTCAGCCT
GGGTCTGTCACTATGGAACATAAAGCTGGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCCTGGGACCAAACAAAT
GCAGATGCTTTTCAGGATACACCGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAAGTCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTGACA
TTGGTTTCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGTCCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAATTAAAAATGTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAC
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAAAATTGAGTTGTATCAAGGAAGTGTATGC
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTTGACTTTGTATGTGAGTCCCTGGTTTTTTTGTATATTGCATCATAG
GACCTCTGGCATTTTGAATTAAGTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTTGATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGAAGTCTTATGATACTTCTTGA
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNRNSKGVCE
ATCEPGCKFGECVGPKNKCRCPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNMCKKAKIKNVTPEPTRTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTGAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGGAATGTCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTCT
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCACAAGATGACAACCAAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAATC
TTTAAGACAAAACATTGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTTTAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRS LCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

05973259-104504

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTCT
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCTACTGTCTGCCACTGCTTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT
CAGTGGTCTCTGGGTTCTCTGCAGCGTGAGGCACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGTCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAAGTGTATCTACAACAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTT
TCCTGGCCCAGAGCCAGAGACCCCGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTCTGGGACGG
GCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGTCTGGGGCACCCTGGTG
CATGAGGTGAGGGGCACATGGTTCTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTACCCGCGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTCGCCGAGGAACCAGAGCCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACCAGCTGCTGAAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCCTGTCTCCCCACCCTGTCTGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTCAGATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTTTTGGGGGGCAGCAGTTTTCTTTTTTTAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSSGGPVLC
LEPDGHVWQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLPDGERG
WVLGRARPGAGISSLQTVPVTLTGPRACSRRLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGCCCCCATTCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTTCGGCAGCGCTGCCGTCAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCCTGA CTGGCAACGTGACCACGTGCGAGATGGCCAACGA
GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGTGCTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCGCAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTACCATCACAGGG
CAGCCTATGACATTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCAAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCTGCCTTATTTACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTC
TCCAATGGCCGTGATACTAGTGATCATGTT CAGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCTTTT
CCCTCCTTCTCCTGCCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGTCTATTCTGGGGCTGGGGCAGTCTTTTCTGGC
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAACCTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSC EEENAGAEDQDGE GEGSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCTGTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACCTTGCTCTGAAGCCCTTGTCGGTAAGCCTTGAACCTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTTCGCTGCTTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCA CTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTTACCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAATTCCTCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTTCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGTATAAG
AAGATGAAGACAAAGATGGGTTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTTTAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTTATTTTTTATACAATGTTCTTTCTTGCTTTGTTTTTTATTTTTTATATTTTTT
CTGACTCCTATTTTAAAGAACCCCTTAGGTTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTGGGAGTTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAGTAGCCGGCGGTGGTGATGCGTGCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCAACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194

><subunit 1 of 1, 211 aa, 1 stop

><MW: 24172, pI: 5.99, NX(S/T): 1

MRLFLWNAVLTLFVTSLLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSL
FHSTHKHNNQPIWFTLGILEALKGWDQGLKGMCMVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCAT
CCAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCTCTCCCGTAGCCCACCCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCTCAACGTCCTCAATGGCTCTGACGCCCCGCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTC
TGAGGAGATGTTCTCCAGTTCGCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTACAACCTGCTACATCATGAACCCCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCCTCCGTGGGGGCTTCTGGCTGTGGTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCTGCTCTCTCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCG
CTCTTGGTGCTTCCCGTGACCTAGGACCCAGGGCCACCTGGGGCCTCCTGAACCCCCG
ACTTCGTATCTCCACCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGG
GGGAGGCAGGAGGCACATGTGAGGGTCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCCGCTGTACCTGCCAGTGCTTGCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCCTGCTCCCAGCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGAGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACCACCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTTCTTTGGGGCTTCAGTGAGAACTCTCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACCTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAACCTGAGATAGTGATAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEEPPEPDRDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

0099889-0000
"55284560"


```

1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <string.h>
4  #include <math.h>
5  #include <time.h>
6  #include <unistd.h>
7  #include <sys/types.h>
8  #include <sys/stat.h>
9  #include <fcntl.h>
10 #include <sys/time.h>
11 #include <sys/mman.h>
12 #include <sys/wait.h>
13 #include <sys/resource.h>
14 #include <sys/param.h>
15 #include <sys/uio.h>
16 #include <sys/xattr.h>
17 #include <sys/zfs.h>
18 #include <sys/dmubuf.h>
19 #include <sys/dmubuf.h>
20 #include <sys/dmubuf.h>
21 #include <sys/dmubuf.h>
22 #include <sys/dmubuf.h>
23 #include <sys/dmubuf.h>
24 #include <sys/dmubuf.h>
25 #include <sys/dmubuf.h>
26 #include <sys/dmubuf.h>
27 #include <sys/dmubuf.h>
28 #include <sys/dmubuf.h>
29 #include <sys/dmubuf.h>
30 #include <sys/dmubuf.h>
31 #include <sys/dmubuf.h>
32 #include <sys/dmubuf.h>
33 #include <sys/dmubuf.h>
34 #include <sys/dmubuf.h>
35 #include <sys/dmubuf.h>
36 #include <sys/dmubuf.h>
37 #include <sys/dmubuf.h>
38 #include <sys/dmubuf.h>
39 #include <sys/dmubuf.h>
40 #include <sys/dmubuf.h>
41 #include <sys/dmubuf.h>
42 #include <sys/dmubuf.h>
43 #include <sys/dmubuf.h>
44 #include <sys/dmubuf.h>
45 #include <sys/dmubuf.h>
46 #include <sys/dmubuf.h>
47 #include <sys/dmubuf.h>
48 #include <sys/dmubuf.h>
49 #include <sys/dmubuf.h>
50 #include <sys/dmubuf.h>
51 #include <sys/dmubuf.h>
52 #include <sys/dmubuf.h>
53 #include <sys/dmubuf.h>
54 #include <sys/dmubuf.h>
55 #include <sys/dmubuf.h>
56 #include <sys/dmubuf.h>
57 #include <sys/dmubuf.h>
58 #include <sys/dmubuf.h>
59 #include <sys/dmubuf.h>
60 #include <sys/dmubuf.h>
61 #include <sys/dmubuf.h>
62 #include <sys/dmubuf.h>
63 #include <sys/dmubuf.h>
64 #include <sys/dmubuf.h>
65 #include <sys/dmubuf.h>
66 #include <sys/dmubuf.h>
67 #include <sys/dmubuf.h>
68 #include <sys/dmubuf.h>
69 #include <sys/dmubuf.h>
70 #include <sys/dmubuf.h>
71 #include <sys/dmubuf.h>
72 #include <sys/dmubuf.h>
73 #include <sys/dmubuf.h>
74 #include <sys/dmubuf.h>
75 #include <sys/dmubuf.h>
76 #include <sys/dmubuf.h>
77 #include <sys/dmubuf.h>
78 #include <sys/dmubuf.h>
79 #include <sys/dmubuf.h>
80 #include <sys/dmubuf.h>
81 #include <sys/dmubuf.h>
82 #include <sys/dmubuf.h>
83 #include <sys/dmubuf.h>
84 #include <sys/dmubuf.h>
85 #include <sys/dmubuf.h>
86 #include <sys/dmubuf.h>
87 #include <sys/dmubuf.h>
88 #include <sys/dmubuf.h>
89 #include <sys/dmubuf.h>
90 #include <sys/dmubuf.h>
91 #include <sys/dmubuf.h>
92 #include <sys/dmubuf.h>
93 #include <sys/dmubuf.h>
94 #include <sys/dmubuf.h>
95 #include <sys/dmubuf.h>
96 #include <sys/dmubuf.h>
97 #include <sys/dmubuf.h>
98 #include <sys/dmubuf.h>
99 #include <sys/dmubuf.h>
100 #include <sys/dmubuf.h>

```

GTTGTATATATGTCCTGAAGTACATCCGTGCATTTTTTTTTTAGCATCCAACCATCCTCCCTTGTA
 GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCACCCNACTAACATCTCAGTCTCTGAA
 AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
 TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC
 TCAATGGCTCTGACGCCCGCCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
 AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCCCTCCAG
 TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGAGTTCTCAGG
 GAACCCAGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
 TTTACAACCTGCTACATCATGAACCCCCC

[illegible]

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCTCAACGTCTCAATGGCTTTGACGCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCC

FIGURE 58

TGCGGCGACCGTCGTACACCAATGGGCCTCCACCTCCGCCCCCTACCGTGTGGGGCTGCTCCCGGATGGCCTCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGGACGTCACCCCCAGTGGTGTGGTCC
CTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCCTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAAGTGTGCTGCCTGTCAATCATGACTGCTGGATTGACAATA
TCAGGCTGGTTTACAACAAAAATCCAGGGCCACCCAGTTTCCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCCTGGACCCCGAGCAAAAGCAGCGTGGGTTCCTATTTCCACACCATGGTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
AAAACGGGCCCTACTTCCTGGCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCGTGGTGC
TGGTTGCCACAGTATGGGCAACATGTACAGCTCTACTTTCTGCAGCGGCAGCCGAGGCCTGGAAGGACAAGT
ATATCCGGGCCCTTCGTGTCACTGGGTGCGCCCTGGGGGGGCGTGGCCAAAGACCTGCGCGTCTGGCTTCAGGAG
ACAACAACCGGATCCAGTCACTCGGGCCCTGAAGATCCGGGAGCAGCAGCGGTGAGCTGTCTCCACAGCTGGC
TGCTGCCCTACAACACTACACATGGTCACTGAGAAGGTGTTCTGTCAGACACCCACAATCAACTACACACTGCGGG
ACTACCGCAAGTTCTTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGAGCTGCACTGCCTCTATGGTACTGGCGTCCCCACACCAGACTCCTTCTACT
ATGAGAGCTTCCCTGACCGTGACCTAAATCTGCTTTGGTGACGCGATGGTACTGTGAAGTGAAGAGTGGCC
TGCAGTGCCAGGCCTGGCAGAGCCGCCAGGAGCACCAAGTGTGTGTCAGGAGCTGCCAGGCAGCGAGCACATCG
AGATGCTGGCCAACGCCACCACCTGGCCTATCTGAAACGTGTGCTCCTTGGGGCCCTGACTCCTGTGCCACAGGA
CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTCAATGGCCACGCGTTTTGCAAAGTTTTGTGA
CTCACCATTCAAGGCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGTGTTTTGTTATCCTTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGGGA
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGAGTGGCTGGGCCCTGGTCCAGTCCCTGCTGGGGCCATG
TGTCCCCCTATTCTGTGGGCTTTTCATACTTGCCTACTGGGCCCTGGCCCCGAGCCTTCCTATGAGGGATGTT
ACTGGGCTGTGGTCTGTACCCAGAGGTCCCAGGGATCGGCTCCTGGCCCCCTCGGGTGACCTTCCCACACACCA
GCCACAGATAGGCCTGCCACTGGTCATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTTCTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGAGTTTGTGCGTTCTTCGTGGTTCCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAAGCTCTGGATTGGGCAGCAGATGTG
CCCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCTTCACCCTGGGACTGTGGTTCCAAGGATGAGAGCAGGGGTGGAGCCATGGCCTTCTGGGAACCTATGGA
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCACCATCA
CACTGCCACCCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCACAGGGCTGAGGATGGGGCTCCTATCCAC
CCTGGCCAGCACCCAGCTTAGTGCTGGGACTAGCCAGAACTTGAATGGGACCTGAGAGAGCCAGGGGTCCCC
TGAGGGCCCCCTAGGGCTTTCTGTCTGCCCCAGGGTGCTCCTATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCTTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGT
GGGGTTCCCAAAGACGCCCTCAGGCTGGACTGAGCTGCTCTCCACAGGGTTTCTGTGCAGCTGGATTTCTCTG
TTGCATACATGCCTGGCATCTGTCTCCCCTTGTTCCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
GATTCTGGCAATAAAAGTACTCTGGATGCTGTAAAAA

09978290 " 104501

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPYRVGLLPDGLLFLLLLLLMLLADPALPAGRHPPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSAVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

09976290 "401501
"09976290"

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTTGAGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAAGTGCCA
GCCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCTCGCTTTTAATGA
CCTCAGCCCCGCCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCCGGGTCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCTCA
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG
CTCCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCAGCTTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCCAGACTCTGTC
TGTGCCGAGTGATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEYSGYNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIQDILLFSLWTFWLVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCCCTGCTCCGAGGCCAGGCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGGATG
CCCGTGGCCGAGGCCCCAGGTGGCTGGCGGGCAGGGGACGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCCCGGGGTACCTCCGCCTGGTGCCCTGTTTGTGCTGCTG
GCCCTGCTCGTGCTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCTTAGGGTACAAGGCGGAGGTGATGGTCAGC
CAGGTGTACTCAGGCAGTCTGCGTGACTCAATCGCCACTTCTCCCAGGATCTTACCCGCCCGGAATCTAGTGCC
TTCCGCAGTGAAACCGCCAAAGCCAGAAGATGCTCAAGGAGCTCATCACCAGCACCCGCTGGGAACCTTACTAC
AACTCCAGCTCCGTCTATTCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTCATTCTCAAATCCCCGAG
CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGACGGCACTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGC
TCGGCTGCCGTCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA
GACATAGCTGCATTGAATTCACGCTGGGTTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCCGGCTG
AAGGGGCCTGACCACCTGGCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGACCTCATGCTCAAACCTCCGG
CTGGAGTGGACGCTGGCAGAGTGGCGGGACCGACTGGCCATGTATGACGTGGCCGGGCCCCCTGGAGAAGAGGCTC
ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTTCTGGCGTCGGGGGCCATCATGGCCGGTC
GTCTGGAAGAAGGGCCTGCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCGTGT
GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGCACCCCGTACTTCCCCAGCTACTAC
TCGCCCCAAACCCACTGTCTCTGGCACCTCAGCGTGCCCTCTCTGGACTACGCTTGGCCCTCTGGTTTGTATGCC
TATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGAGCATCCAGAACAGGAGGCTGTGT
GGCTTGCGCATCCCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTACC
TCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACAGTCGGACCCCTGCCCTCGGA
GAGTTCTCTGTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGACTGCCCCAACGGCCTGGAT
GAGAGAACTGCGTTTGAGAGCCACATTCCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCAAGGTCTGT
GATGGGCAGCCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACC
TTCCAGTGTGAGGACCGGAGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGCCCCGACTGCAGGGACGGC
TCGGATGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCTCCAGCCGCAATTGTTGGTGGAGCTGTGTCTCCGAG
GGTGAAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGTGACACATCTGTGGGGGGCCCTCATCGCTGACCGC
TGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCTCCACGGTGTGTGGACCGTGTCTCTGGGC
AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCAC
GAAGAGGACAGCCATGACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCCGCGTGGC
CCCGTCTGCCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCTGCACTGCTGGATTACGGGCTGGGGCGCCTTG
CGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCCAAGGACCTGTGCAGCGAG
GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGT
GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTG
GGCTGTGGCCGGCCTAACTACTTCCGCGTCTACACCCGCATCAGAGTGTGATCAGCTGGATCCAGCAAGTGGTG
ACCTGAGGAAGTGCCTCCCTGCAAAGCAGGGCCACCTCTGGACTCAGAGAGCCAGGGCAACTGCCAAGCAGG
GGGACAAGTATTTCTGGCGGGGGGTGGGGGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTTGTCTCGTCC
CTGATGTCTGCTCCAGTGATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAAGAGCTCCCTGAGGACC
CAGGCCCCACCCAGCCCTTCTGCCTCCCAATTCTCTCTCCTCCGTCCCCTTCTCCACTGCTGCCTAATGCAAG
GCAGTGGCTCAGCAGCAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG
CTGTTTGGGCAGCCTTGCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCTGGTCTAATTGGGATCTGGGAAT
GGAAGGTGCTCCCATCGAGGGGACCCCTCAGAGCCCTGGAGACTGCCAGGTGGGCCTGCTGCCACTGTAAGCCAA
AAGGTGGGAAGTCTGACTCAGGGTCTTGCCCCACCCCTGCCTGCCACCTGGGCCCTCACAGCCAGACCCCT
CACTGGGAGGTGAGCTCAGCTGCCCTTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152

><subunit 1 of 1, 802 aa, 1 stop

><MW: 88846, pI: 6.41, NX(S/T): 7

MPVAEAPQVAGGQGDGGDGEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YNNSSSVYSFGEGPLTCFFWFILQIPEHRRMLLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCRYRYSYVGQGVRLRLKGPDLASSCLWHLQGPDLML
KLRLEWTLAECDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLLDSQGVLTSTPYFPSYYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDDGQPDCLNGSDEEQCQEGVPCGTFTTFQCEDRSCVKKPNPQCDGRPDGRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITAAHCFQEDSMASVTL
WTVFLGKQVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSSGGLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTTCGGGGTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCTTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGCCGTGCGCCC
CGTCTGCCTGCCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCACAGGACCTGTGCAGCGAGGCCATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGTGA CTGAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACTGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGG
CTCCGTGCCGCCAAGTTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAAA
ACTTCTGTTTTCTTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCCCTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTGCGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTTGAAAACCTGGCTGCACAAACTCGTGTCAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG
TGGATGTCCTCCACACCTACACGCGTTCCTTCGGCTTGAGCATTGGTATTAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTTGCCTTCCAGTGCACT
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTTCAGAGGTAACCTTCAGTCCCTGGAGTGTCCTTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL
PLAHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIAKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLEFVDSLNVQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCCGGGCCGAGCCACCTCTTCCCCTCCCCCGC
TTCCCTGTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT
CGGCAAAGTTTGGCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGCGCTCG
CTGCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
AGAAGAGTGC GGCGGCGGACGGAGAAAACTCCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCG
CCGCTCCCCCGCCCCAGCCCTGGCATCCAGAGTACGGGTCGAGCCCGGGCCATGGAGCCCCCTGGGGAGGCGG
CACCAGGGAGCCTGGGCGCCCCGGGCTCCGCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTCCG
GCACCTCTGGACAGCCCAGGATGCTGTTGGCCACCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG
ACCGGATTATTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTAGAAGTGCAGGGCACCTTACAGA
GGCCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAACAGACTG
TCACCATCAGGTTCAGAAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCTACGCTCCCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCAGCCCTCTGCAGCTGCCCGGGGCAACGTCAACATCACTTACAGCTATG
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG
AGTTTCAGTGCCTGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT
CTGATGAAGCAGGTTCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCCCCTCCCTGCCTTGCAATG
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCCTCCTGGATATACACACCTAGCCTCAGTCTCCCACCCCCAGT
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTCACAGCCCTGGACTTGGGCTTTG
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCTTGAGAGCTCCCGACTACTGCGTAGTCTACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTACACACAGTTGCTTGGAGCA
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCCTGTGGCTTAGGCT
CTGGCCTGGGAGCTGGCGAAGGCTTAGGTGAGCGCTGCTACAGTGAAGCACAGCGCTGTGACGGCTCATGGGACT
GTGCTGACGGCACAGATGAGGAGGACTGCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT
CTGGTGCCACAGCTGCTACCTGCCTGCTGACCGCTGCAACTACCAGACTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTCCGCATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCCTGTATGAGACGTGGGTGTGCG
ATGGGCAGCCAGACTGTGCGGACGGCAGTGTGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
TTTCACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGGATGGAGGCTGAGATTGTGAGCAGCAGGCACCCCC
CTTCCTACGGGCAGCTCATTGCCAGGGTGCCATCCCACCTGTAGAAGACTTTCCTACAGAGAATCCTAATGATA
ACTCAGTGTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
CCCGCCGTGCTCAGCGGGGCCGCTTGATGCGACGCTGGTACGCCGTCTCCGCCGCTGGGGCTTGCTCCCTCGAA
CCAACACCCCGGCTCGGGCTCTGAGGCCAGATCCAGGTACACCTTCTGCTGCTCCCCTTGAGGCCCTAGATG
GTGGCACAGGTCCAGCCCGTGAGGGCGGGGCGAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCCATCA
AGGCTCCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCTGTGTCAGCCTGGGGCCCC
CAGGACCAACCCGGAGCCCCCTGGACCCCAACAGCAGTCTTGCCCTGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG
CTCTACTGAGGCCTCTCCCCTGGGGGCTCTACTCATAGTGGCACAACCTTTTAGAGGTGGGTGAGCCTCCCCCTCC
ACCACTTCCTTCCCTGTCCCTGGATTTTCAAGGACTTGGTGGGCTCCCGTTGACCCCTATGTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCAGTCCCT
TCACCACCACCTGCTCCCCACGCCACCACCTTTGGGTGGCTGTTTTTAAAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAACCTAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGTTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG
CTCACA AAAAGAGTGCACA AATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA
GGAATCATACATCTC

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPDRIIFPNHACEDPPAVLLEVOGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCSVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP
VPSLPCNVTLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLRLSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDGQPD CADGSDEWDCS
YVLPRKVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGRLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLP SLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTATTTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAAGAAGTTTTGTAATTTTATATTACTTTTTAGTTTGATACTAAGT
ATTAAACATATTTCTGTATTCTTCAAAAAAAAAAAAAAAAAAAAA

657249-2003-0001

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSM TFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCGTGGGCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCAG
CTGGCGTGCTGGACGAATTTCGGAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACG
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTTCGGCATTTCCTCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCTCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATTGAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCCACCCGTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIIVDSGTTLLR
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFRCL
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCGCTGGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG
TCACGGGCGCCAAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAAACCTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTGCGGGACGTCTTGAAGCTCAAACGCCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTGCGCCATGTTCTTGATGGCTGCGCCCACTTTTG
CGCCCATTTGGCTTGGCTGGTGTCCGGGCACCAAGAGGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCACCCTGAGGAGCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCTTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTTCAGGCCCCACCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAAGTATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVVGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLLLTHLLLPCLKACAPSRVVVVASAAHCRGRLDKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFRLRHVPGWLRPLLR
PLAWLVLRAPRGGAAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDRAAHLWEASKRL
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPQSSPDLSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

"09701" 65234650

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATCATGAGGAGGCATGGCTCAG
GACTCCCCGCCCCAGATCCTAGTCCACCCCCAGGACCAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
CAAGCCTCAGGCCAGCCACCTCCCACCATCCGCTGGTTGTCTGAATGGGCAGCCCCTGAGCATGGTGCCCCAGAC
CCACACCACCTCCTGCCTGATGGGACCTTCTGCTGCTACAGCCCCCTGCCCGGGGACATGCCACGATGGCCAG
GCCCTGTCCACAGACCTGGGTGTCTACACATGTGAGGCCAGCAACCGGCTTGGCACGGCAGTCAGCAGAGGCGCT
CGGCTGTCTGTGGCTGTCTCCTCCGGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGGTGAGCAG
TTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCACAGTCTCATGGTGGAAGATGGGAAACCCCTG
GCCCTCCAGCCCGGAAGGCACACAGTGTCCGGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGG
ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGGTTTCCATCCAGGAGCCC
CAGGACTACACGAGCCTGTGGAGCTTCTGGCTGTGCGAATTCCAGCTGGAAATGTGACACTGCTGAACCCGGAT
CCTGCAGAGGGCCCCAAGCCTAGACCGGCGGTGTGGCTCAGCTGGAAGGTCACTGGCCCTGCTGCGCCTGCCCAA
TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCCGGGAGGCCAGGGAGCTCCGTGGGCAGAGGAGCTGCTGGCC
GGCTGGCAGAGCGCAGAGCTTGGAGGCCTCCACTGGGGCCAGACTACGAGTTCAAAGTGAGACCATCCTCTGGC
CGGGCTCGAGGCCCTGACAGCAACGTGTCTCTGAGGCTGCCGGAAGAGTGCCAGTGCCCACTCAGGAA
GTGACTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGGTCCACCACCTGCTGAAAACCAATGGCATC
ATCCGTGGCTACCAGTCTGGAGCCTGGGCAACACATCACTGCCACCAGCCAAGTGGACTGTAGTTGGTGAGCAG
ACCCAGCTGGAAATCGCCACCCATATGCCAGGCTCCTACTGCGTGCAAGTGGCTGCAGTCACTGGTGTGGAGCT
AGGGAGCCCCAGTAGACCTGTCTGCCTCCTTTTAGAGCAGGCCATGGAGCGAGCCACCAAGAACCCAGTGAGCAT
GGTCCCTGGACCCCTGGAGCAGCTGAGGGCTACCTTGAAGCGGCCCTGAGGTCACTGCCACCTGCGGTGTTCACCTC
TGGCTGTCTCTTCTGGGCACCGCGTGTGTATCCACCGCCGGCGCCGAGCTAGGGTGACCTGGGCCCAGGTCTG
TACAGATATACCAAGTGGAGTGCCATCCTAAAACACAGGATGGATCACAGTGAATCCAGTGGTTGGCAGACACT
TGGCGTTCCACCTCTGGCTCTCGGGACCTGAGCAGCAGCAGCAGCCTCAGCAGTCCGGCTGGGGGCGGATGCCCGG
GACCCACTAGACTGTCTGCTCCTTGTCTCTTGGGACTCCCGAAGCCCCGGCGTGCCCTGCTTCCAGACACC
AGCACTTTTATGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCAGCCAGGCCAAGTCCCCAGGTCCCAGCT
GTCAGGCGCCTCCACCCAGCTGGCCAGCTCTCCAGCCCTGTTCCAGCTCAGACAGCCTCTGCAGCCGAGG
GGACTCTCTTCTCCCCGCTTGTCTCTGGCCCCCTGCAGAGGCTTGAAGGCCAAAAGAAGCAGGAGCTGCAGCAT
GCCAACAGTTCCCCACTGCTCCGGGGCAGCCACTCCTTGGAGTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
AAGAACCTTTCCCAAAGCCCAGGAGCTGTGCCCCAAGCTCTGGTTGCCTGGCGGGCCCTGGGACCGAAACTCCTC
AGCTCCTCAAATGAGCTGGTTACTCGTCACTCTCCCTCCAGCACCCTCTTCTCATGAAACTCCCCAACTCAG
AGTCAACAGACCCAGCCTCCGGTGGCACCAAGGCTCCCTCCATCCCTGCTGCCAGCAGCCCCCATCCCCATC
CTTAGCCCCCTGCAGTCCCCCTAGCCCCAGGCCTCTTCCCTCTCTGGCCCCAGCCAGCTTCCAGTCCGCTGTCC
AGCTCCTCACTGTCTATCCCTGGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
CTCAGTGAGGGTGAGGAGACTCCCAGGAACAGCGTCTCTCCATGCCAAGGGCTCCTTACCCCCCACCACCTAT
GGGTACATCAGCGTCCCAACAGCCTCAGAGTTACGGACATGGGCAGGACTGGAGGAGGGGTGGGGCCCAAGGGG
GGAGTCTTGCTGTGCCACCTCCGGCCCTGCCTCACCCCCACCCAGCGAGGGCTCCTTAGCCAATGGTTGGGGC
TCAGCCTCTGAGGACAATGCCGCCAGCGCCAGAGCCAGCCTTGTGAGTCTCTCCGATGGCTCCTTCTCGCTGAT
GCTCACTTTGCCCGGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCCGTCTAGAGCCCAGGGAGGCAGACTGC
GTCTTCATAGATGCCTCATCACCTCCCTCCCCACGGGATGAGATCTTCTGACCCCCAACCTCTCCCTGCCCTG
TGGGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTGAGCCACACCCAGCGCTGGGAAGGGGGATGCCCTCC
TGGCCCCCTGACTCTCAGATCTCTTCCCAGAGAAGTCAGCTCCACTGTCTATGCCCAAGGCTGGTGTCTCTCT
GTAGATTACTCTGAACCGTGTCTCTGAGACTTCCAGACGGGAATCAGAACCCTTCTCTGTCCACCCACAAG
ACCTGGGCTGTGGTGTGTGGGTCTTGGCCTGTGTTTCTCTGCAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAG
TTCTCCCTCCACGATTGTGAAAAACAAATGAAAAACAAATTAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAA
ACATCATCTCCACCTGACTCCTAGCCACTGCTTTCTCCTCTGTGCCATCCACTCCACCACCAGGTTGTTTTGGC
CTGAGGAGCAGCCCTGCCTGCTGCTCTTCCCCCACCATTGGATCACAGGAAGTGGAGGAGCCAGAGGTGCCCTT
GTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGGAGCCCCCTCTCAGCCTTACCT
GGGCCCCCTCTCTAGAGAAGAGCTCAACTCTCTCCCAACCTCACCATGGAAAGAAAAATAATTATGAATGCCACTG
AGGCACTGAGGCCCTACCTCATGCCAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
TATGAGACCGTAGGTCAAAGACCATCCTCGTACTGTTGTCACTATGAGCTTAAGAAATTTGATACCATAAAAT
GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGPGRMSCQASGQPPPTIRWLLNGQPLSMVPPDPHHLLP
DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGWP
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLLPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA
WKAKKKQELQHANSPLLGRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPDPDSQISSQRSQLHCRMPKAGASPDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

09759-10101

FIGURE 78

CTCCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
 CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
 CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCCTATATGCAGAAGAAGAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
 ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTCTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCTTCTCCAC
 CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAAACCCAGCCCC
 CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGCCCCCTCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCATGACAGC
 TGGACTCCACCTCAGCAGAGGACACCACTCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGCTGCTGAGCCTTCTGTGAGC
 CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
 CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
 CCTTCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCACCTCAGCCTCAGAG
 TCCAGCTGCCCGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGCACTGTTCCA
 GCCTGACCTAGAAGCGTTTGTAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCGGGC
 CTCATGCCCAGTGTGCGGACCCTGCCTTCTCTCCACTCCAGACCCCACTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCAGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
 GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGACG
 GACTCTGAATTCTAACAATGCCCAGTGACTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
 TCCAGGCCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCAGGACCGGCACAGAAGTGG
 TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCCTTTGGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCCTATGGGTTCTGGTGGCTAGAGA
 GAAAAGTAGAAAACAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
 GCACAACTACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
 GCTGCCCAGGCTGGAGTGCAGTGGCAGCATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT
 CTTCTGCCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCAACACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCCTGAC
 CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTTAAAAAGTGAAATTAAGAGTTGTTTCAATATGCAAACTTGGAAG
 ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTACCCATAGTCTACCCAGAGACTATCAT
 TATTTGTTTTGTTGTACTTCTTCCACTCTTTTCTTCTTACATAATTGCGCGGTGTTCTT
 TTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
 ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGP EEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSIRDSRQELSLIVTLWNLT LQDAGEYWCGVEKRGPD E SLLISLFV
FPGPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSP PAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRI
LAPVLVLLSLLSAAGLIAFC SHLLLWRKEAQQATETQRNEKFWLSRLTAE EKEAPSQAPEGD
VISMPPLHTSEEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

[illegible]

TTGTGACTAAAAGCTGGCCTTAGCCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTA CTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCC CAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTTACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCNLTPNVPNVCRMYSDDLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homolgous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

09978299-101501
T0101 6692660

```

length = 1000000
seed = 123456789
set.seed(seed)
n = 1000000
n2 = 1000000
n3 = 1000000
n4 = 1000000
n5 = 1000000
n6 = 1000000
n7 = 1000000
n8 = 1000000
n9 = 1000000
n10 = 1000000
n11 = 1000000
n12 = 1000000
n13 = 1000000
n14 = 1000000
n15 = 1000000
n16 = 1000000
n17 = 1000000
n18 = 1000000
n19 = 1000000
n20 = 1000000
n21 = 1000000
n22 = 1000000
n23 = 1000000
n24 = 1000000
n25 = 1000000
n26 = 1000000
n27 = 1000000
n28 = 1000000
n29 = 1000000
n30 = 1000000
n31 = 1000000
n32 = 1000000
n33 = 1000000
n34 = 1000000
n35 = 1000000
n36 = 1000000
n37 = 1000000
n38 = 1000000
n39 = 1000000
n40 = 1000000
n41 = 1000000
n42 = 1000000
n43 = 1000000
n44 = 1000000
n45 = 1000000
n46 = 1000000
n47 = 1000000
n48 = 1000000
n49 = 1000000
n50 = 1000000
n51 = 1000000
n52 = 1000000
n53 = 1000000
n54 = 1000000
n55 = 1000000
n56 = 1000000
n57 = 1000000
n58 = 1000000
n59 = 1000000
n60 = 1000000
n61 = 1000000
n62 = 1000000
n63 = 1000000
n64 = 1000000
n65 = 1000000
n66 = 1000000
n67 = 1000000
n68 = 1000000
n69 = 1000000
n70 = 1000000
n71 = 1000000
n72 = 1000000
n73 = 1000000
n74 = 1000000
n75 = 1000000
n76 = 1000000
n77 = 1000000
n78 = 1000000
n79 = 1000000
n80 = 1000000
n81 = 1000000
n82 = 1000000
n83 = 1000000
n84 = 1000000
n85 = 1000000
n86 = 1000000
n87 = 1000000
n88 = 1000000
n89 = 1000000
n90 = 1000000
n91 = 1000000
n92 = 1000000
n93 = 1000000
n94 = 1000000
n95 = 1000000
n96 = 1000000
n97 = 1000000
n98 = 1000000
n99 = 1000000
n100 = 1000000

```

AGCCGCTGCCCCGGGCGGGCGCCCGCGGGCACCATGAGTCCCCGCTCGTGCCTGCGTTC
GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTCTGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAAGTGTCTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCCTTCGT
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGACAGTGGGGAGCT
GGAGAAGTGC GGCTGTGACAGGACAGTGCATGGGGTCAGCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCACGGGGTGTCAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGCTGGGCACGAGGGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCTACTGGTGTCTTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATTGACCGCCTGCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTTTGTCTGGTTTGGTTTTTGGGTCCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCCGATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTGATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCTCTGGAAAAAAGAACTCTTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATAACAGGTCAAGGGCACCAGGTTCAATTCAGCCCTTACATGGACAGCTAGA
GGTTCGATATCTGTGGGTCTCTTCCAGGCAAGAAGAGGGGAGATGAGAGCAAGAGACGACTGAA
GTCCCACCCTAGAAACCCAGCCTGCCCCAGCCTGCCCCTGGGAAGAGGAAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCGAGTCACCTTTCACAGCGCTGTTCTCTCATGAACTGAAAAACACACAC
ACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCCTTTGCAGTCATGCCCGAGTCACCTTTCACAGCACTGTTCTCTC

[illegible]

<subunit 1 of 1, 351 aa, 1 stop

MSPRSCLRSLRLLVFAVFSAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRWNCSTLDSL PVFGKVVTQGTREAAFYAIISSAGVAFAV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSEVKT CWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCFKHWCCFVKCROCRLVELHTCR

Signal peptide:

N-glycosylation sites.

Wnt-1 family signature.

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTC
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAAGTTCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCG
TGCCCCACGTTCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTATGATGC
TGCAACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAAGTGCTGACCCCGCCAGTGCCCTGGAGCCGCGCCCATTCAGCATGTCTGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAAGTGA
AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSSEEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTOGLAEAGRGREDVRTELFRALEAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEPNDAWGRENVCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCCTTCTCCTCCTTTCTCCACGTCTTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACTGGAGCCTCATTTGGCCGGCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGAAAAACCCAGCCCGGCCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCAAGCAGTACCCCTGTTCGCCCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTACGTAAACGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTTCGGCGCCCGCCGTCCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTCTGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCCTCTCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCTGCCCAGC
AGGGACAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCCCAACAACGGGAGCCCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAACTGCGTCTAAGACCAGAGCCCCGAGCCCCCTGGGGCCCCCG
GAGCCATGGGGTGTGCGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCTTCATCG
TCCAGGGGCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAEERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFVGVDSDLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAPVLP SRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGR LGTKSRTRYVRVQPA
NNGSPCPELEEEAECVPDNCV

Important features:

Signal peptide:

amino acids 1-26

"105707" "663660"

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACCTTTGTTTGAAGCTTATTATGACAATACCATTTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTTCATTCACGGTTGCGTTTAAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTCGAGCAGATGAACT
TAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT
GAGGTTTTGTTTAAATCCTTTTGATGACATCATTCGAAGGGAAATTAAGGCTGAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCGAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAACGGGAACTCTTAGCAG
CAAAACAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTTTTGCAAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCTT
TCCACAAAAA

FIGURE 89

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919

><subunit 1 of 1, 472 aa, 1 stop

><MW: 53847, pI: 5.75, NX(S/T): 2

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEAEEEEEEVN RVVSQSMKGKSKSSHDLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEEAPPDGAVAEYRREKQKYEALRK
QQSKKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQSDTFEIYDPRNPVNKRREESKKLMREKKERR

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCAGACCCATTTTCGCCCTTGCTGACGGCGTCCG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTGGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTTGACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGACCTCCAGAACCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTTGGGGCCAGCTGGCCTCG
ATGTACGTGACGACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
GAGAAAGTGGTGTGTTTTCTGCAGCAGGCTCGGCATCTTCGGCTTCTGAGCACGGA
CGACAGCCACGCGCGCGGGAACCTGGGGGCTGCTGGACAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCAGGAAATGTGACCCTGTTCCGCCAGTCGGCG
GGGGCCATGAGCATCTCAGGATGATGATGTCAACCCTAGCCTCGGGTCTCTTCCATCGGGC
CATTTCCAGAGTGGCACCGCTTATTGAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGTTGCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCCTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTCCAAC
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCAGATGACCTTTTGGTGTCTCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCCCTTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCAGTTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCCATGATGCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTCCCTTCTTCAAATCCT
CCCACCCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTACCATCCATCCTGCTCAACCTTGTTCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTGTGACATGGAGCAAACTTTTGGTAGTTTGGGA
TCTTCTCTCCACCCACACTTATCTCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTGATTAGAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTQKGQMHVGKTPIQVFLGVPPFSRPP
LGILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNNKMRFLQLNFQRPDEEIIWSMSPVVDGVVIPPDD
PLVLLTQGVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDKMLRNRMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

[illegible]

GAGAACACAGGCTGTCTCAGGCAGGCGCTGCGCCTCCTATGCGGAGATGCTACTGCCACTGCT
 GCTGTCCTCGCTGCTGGGCGGGTCCCAGGCTATGGATGGGAGATTCTGGATACGAGTGCAGG
 AGTCAGTGATGGTGCCGGAGGGCCTGTGCATCTCTGTGCCCTGCTCTTTCTCCTACCCCCGA
 CAAGACTGGACAGGGTCTACCCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
 CAAGGGTGCTCCTGTGGCCACAAACCACCAGAGTCGAGAGGTGGAAATGAGCACCCGGGGCC
 GATTCAGCTCACTGGGGATCCCGCCAAGGGGAACTGCTCCTTGGTGATCAGAGACGCGCAG
 ATGCAGGATGAGTCACAGTACTTCTTTTCGGGTGGAGAGAGGAAGCTATGTGACATATAATTT
 CATGAACGATGGGTTCTTTCTAAAAGTAAAGTGCTCAGCTTCACGCCCAGACCCAGGACC
 ACAACACCGACCTCACCTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
 GTCCGACTCCGTGTGGCCTATGCCCCCAGAGACCTTGTTATCAGCATTTTACGTGACAACAC
 GCCAGCCCTGGAGCCCAGCCCCAGGGAAATGTCCCATACCTGGAAGCCCCAAAAGGCCAGT
 TCCTGCGGCTCCTCTGTGCTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCTGCGAG
 AACAGAGTCCTCTCCTCGTCCCATCCCTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCGG
 GGTGAAGGCTGGGGATTACAGGCGCTACACCTGCCGAGCGGAGAACAGGCTTGGCTCCCAGC
 AGCGAGCCCTGGACCTCTCTGTGCGATATCCTCCAGAGAACCTGAGAGTGATGGTTTCCCA
 GCAAACAGGACAGTCTTGAAAACCTTGGGAACCGCACGTCTCTCCAGTACTGGAGGGCCA
 AAGCCTGTGCCTGGTCTGTGTACACACAGCAGCCCCCAGCCAGGCTGAGCTGGACCCAGA
 GGGGACAGGTTCTGAGCCCCCTCCAGCCCTCAGACCCCGGGGTCTTGAGCTGCCTCGGGTT
 CAAGTGAGACGAAAGGAGAGTTACCTGCCACGCTCGGCACCCACTGGGCTCCCAGCACGT
 CTCTCTCAGCCTCTCCGTGCACTATAAGAAGGGACTCATCTCAAAGGCATTCTCCAACGGAG
 CGTTTCTGGGAATCGGCATCACGGCTCTTCTTTTCTCTGCTGGCCTGATCATCATGAAG
 ATTCTACCGAAGAGACGGACTCAGACAGAAACCCCGAGGCCAGGTTCTCCCGGCACAGCAC
 GATCCTGGATTACATCAATGTGGTCCCGACGGCTGGCCCCCTGGCTCAGAAACGGAATCAGA
 AAGCCACACCAAACAGTCCTCGGACCCCTCCTCCACCAGGTGCTCCCTCCCAGAATCAAAG
 AAGAACAGAAAAAGCAGTATCAGTTGCCAGTTTCCCAGAACCCTAAATCATCCACTCAAGC
 CCCAGAATCCAGGAGAGCCAAAGAGGAGCTCCATTATGCCACGCTCAACTTCCCAGGCGTCA
 GACCCAGGCTGAGGCCCGGATGCCCAAGGGCACCCAGGCGGATTATGCAGAAAGTCAAGTTC
 CAATGAGGGTCTCTTAGGCTTTAGGACTGGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
 GTTGAAAGATAACAGAGTGCAAAGTTTCTTTCTCTCCCTCTCTCTCTCTCTTTCTCTCTCT
 CTCTCTTTCTCTCTCTTTTAAAAAACATCTGGCCAGGGCACAGTGGCTCACGCCTGTAATC
 CCAGCACTTTGGGAGGTTGAGGTGGGCAGATCGCCTGAGGTGCGGGAGTTGAGACCAGCCTG
 GCCAACTTGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCAGG
 CGCCTGTAATCCTACCTACTTGGGAAGCTGAGGCAGGAGAATCACTTGAACCTGGGAGACGG
 AGGTTGCAGTGAGCCAAGATCACACCATTGCACGCCAGCCTGGGCAACAAAGCGAGACTCCA
 TCTCAAAAAAAAAAATCCTCCAAATGGGTGGGTGTCTGTAATCCAGCACTTTGGGAGGCTA
 AGGTGGGTGGATTGCTTGAGCCAGGAGTTGAGACCAGCCTGGGCAACATGGTGAAACCCC
 ATCTCTACAAAAAATACAAAACATAGCTGGGCTTGGTGGTGTGTGCCTGTAGTCCAGCTGT
 CAGACATTTAAACCAGAGCAACTCCATCTGGAATAGGAGCTGAATAAAATGAGGCTGAGACC
 TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCCG
 TACAAGATACAGGTATAAAGACTTTGCTGATAAAACAGATTGCAGTAAAGAAGCCAACCA
 ATCCCACCAAAACCAAGTTGGCCACGAGAGTGACCTCTGGTCGTCTCACTGCTACACTCCT
 GACAGCACCATGACAGTTTACAAATGCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
 AAAGGGGGAGGAATGAATAATCCACCCCTTGTTTAGCAAATAAGCAAGAAATAACCATAAAA
 GTGGGCAACCAGCAGCTCTAGGCGCTGCTCTTGTCTATGGAGTAGCCATTCTTTTGTTCCTT
 TACTTTCTTAATAAACTTGCTTTCACCTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPKGNCSLVIRDAQMQUESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSTPRPDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACCTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCATCTCATTTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCTTTCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCTATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAACCTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTTGAAGAGGAT
GATTGCGCAAACATTTGACATTATTACTAATCATGAAACTCAAGAAGGAATAGTTATATTTAA
AAAGAAAGTGGATTGTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCCCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTTCATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCTCCTTCCATATTATGTATTTGAAGTTTTTGA
AGAAACCCACAGGGATCATTGTTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCAATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAAGAGCACCATTT
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCATTTTGACTAATAGAACTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAAATGGAATCCCGTCACTTACAAGTACAAACACCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTGCATTATGATCATA
TTTGGGTTTATTTTTTTGACTTTGGGTTTAAAAACAACGGAGAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTTTCAAGAGAGAATATATTCCAATATGATGATGAAGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCAGTCAAATAATTAGG
GCTTTTTTACCATCAAATTTTTTAAAAGTGCTAATGTGTATTGGAACCAATGGTAGTCTTAA
AGAGTTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTTAAGCTACCTACATGCTGTATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAATAATG
TAGGAAGATATTTAAAAGTAGATGAGAGGACA CAAGATGTAGTCGATCCTTATGCGATTATAT
CATTATTTACTTAGGAAAGAGTAAAAATACCAAACGAGAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTGTTTCCTTTAT
TTAAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVS DINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTG VIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLS D
VNDNKPIFKESLYRLTVSESAPTGT SIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTNR TG FNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCD CGDSGSTQTCQYQELVLSMGFKTEV IIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPD SAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGS LAGSLSSLES AVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

[illegible]

ATTTC AAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAACTTT CAGATGTTAATGACAA
TAAGCCTATATTTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTGCGAAACATTTGACATTATT

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGCGCGCAGCGAGCAGGGCTCCCCGCCTTAACCTTCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTGCCCACCTGCAAACCTCTCCGCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTCAATGGCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCTTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTCTTGACAGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAGAATTCTATGACCCTATGACCCACAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCCTGTCCCCGAAAAACAACCTTTACCCAAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTGTGACACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAAA
ACCATGTGTTAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCCTCA
ATATAGGAGGGAAGATTTTCCATTGTATTACTGCTTCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTA
ATAGACAGTAAATACTATTCTCATTATGTTGATACTAGCATACTTAAATATCTCTAAAT
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTCGTCC
TTATATACATATGAACAGTCAAATCATTACTTCTTCTTCAATTAGCTTTGGGTGCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGATTTATTACAAATCAGAACTTTGGAGGCAAAATCTTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
AACACAACCTTTATTGATTGAATTTTAAAGCTACTTATTCATAGTTTTATATCCCCCTAACT
ACCTTTTTGTTCCTTCAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTTA
TATCTTCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACATATGCCTATGTAGTGTCTTTATTTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACTTCAT
GTGGTTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCATTCTTTCACTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAACAACCTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGGCAAGGCATTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAACCTGCATACGTTTTGGTGTGCTTTTCAAATGTTTGA
AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI~~L~~AF~~L~~GWIGAI~~V~~STALPQWRIYSYAGDNIVTAQAMYEG~~L~~WMSCV~~S~~QSTGQI
QCKVFDSL~~L~~NLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDE~~V~~QKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMT~~P~~VNARYEFGQALFTGWAAASLCLLGGALLCCSC
PRKTT~~S~~YPTPRPYPKPAPSSGKD~~Y~~V

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

09578399-101501

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTTCGAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT
GTTCCCTGTCCC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

ACCCTTGACCCAACGCGGCCCGGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
 GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCATCNTCAGCACTGCCCTGCCCAGTG
 GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
 GGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
 GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
 TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
 GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGGATATTTCTTCTTGACGGTCTGGCTA
 TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

[illegible]

GGGCCCGACCATTTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANCAACGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTG CAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

09978269 "104504
"09978269 "104504

[illegible]

[illegible]

TCATAGGGGGGCGGATATTTTTCTTGACAGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCTATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCTTG

FIGURE 106

TTCCTGGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGGCTGTGGATGTCNGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTGCGAA

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCCGCGCGTCCCCG
ACCGGTCCCCGCTTTTTTGTAAACCTTAAAGCGGGCGCAGCATTAACGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCACTCTATGTTTGCTCCAACCTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAAACCTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACACATGATGTAGAAATAAATAAAATTATATCCACAACCTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTACAGGCTACGGGAGGAGAACAAAGCAGTTCAAG
GAAGAAGATGGAGTGCAGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTTCAGCATTAGC
CCAACCTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCCTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTTGGTAGGCCCTTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCCTTCTCGCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCACGCAT
GCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCAAGTGGCCCACTCCCGGCCAGGCTG
CTTTCCTGTCTTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTAAGTGCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTGTTATTCAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGTTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTTCTGGGTGAGTAAATAACAACCTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCCCCAAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMQLQPFDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

09739-4304
"0628450"

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTTGTCAATTTTGTACATTTGCTCTATGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTAACTTTAAAATGAGC

09978299-101501
"09978299-101501"

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTA CTGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAAC TGTATTTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

0007229-101201

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

"0929-101501"

```

% test 1: 1000000000
% test 2: 1000000000
% test 3: 1000000000
% test 4: 1000000000
% test 5: 1000000000
% test 6: 1000000000
% test 7: 1000000000
% test 8: 1000000000
% test 9: 1000000000
% test 10: 1000000000
% test 11: 1000000000
% test 12: 1000000000
% test 13: 1000000000
% test 14: 1000000000
% test 15: 1000000000
% test 16: 1000000000
% test 17: 1000000000
% test 18: 1000000000
% test 19: 1000000000
% test 20: 1000000000
% test 21: 1000000000
% test 22: 1000000000
% test 23: 1000000000
% test 24: 1000000000
% test 25: 1000000000
% test 26: 1000000000
% test 27: 1000000000
% test 28: 1000000000
% test 29: 1000000000
% test 30: 1000000000
% test 31: 1000000000
% test 32: 1000000000
% test 33: 1000000000
% test 34: 1000000000
% test 35: 1000000000
% test 36: 1000000000
% test 37: 1000000000
% test 38: 1000000000
% test 39: 1000000000
% test 40: 1000000000
% test 41: 1000000000
% test 42: 1000000000
% test 43: 1000000000
% test 44: 1000000000
% test 45: 1000000000
% test 46: 1000000000
% test 47: 1000000000
% test 48: 1000000000
% test 49: 1000000000
% test 50: 1000000000
% test 51: 1000000000
% test 52: 1000000000
% test 53: 1000000000
% test 54: 1000000000
% test 55: 1000000000
% test 56: 1000000000
% test 57: 1000000000
% test 58: 1000000000
% test 59: 1000000000
% test 60: 1000000000
% test 61: 1000000000
% test 62: 1000000000
% test 63: 1000000000
% test 64: 1000000000
% test 65: 1000000000
% test 66: 1000000000
% test 67: 1000000000
% test 68: 1000000000
% test 69: 1000000000
% test 70: 1000000000
% test 71: 1000000000
% test 72: 1000000000
% test 73: 1000000000
% test 74: 1000000000
% test 75: 1000000000
% test 76: 1000000000
% test 77: 1000000000
% test 78: 1000000000
% test 79: 1000000000
% test 80: 1000000000
% test 81: 1000000000
% test 82: 1000000000
% test 83: 1000000000
% test 84: 1000000000
% test 85: 1000000000
% test 86: 1000000000
% test 87: 1000000000
% test 88: 1000000000
% test 89: 1000000000
% test 90: 1000000000
% test 91: 1000000000
% test 92: 1000000000
% test 93: 1000000000
% test 94: 1000000000
% test 95: 1000000000
% test 96: 1000000000
% test 97: 1000000000
% test 98: 1000000000
% test 99: 1000000000
% test 100: 1000000000

```

GGTGGCCCATTCCCG GCCCAGGCTGCTTTCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAAGT
TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACGTATTATTCAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

097339-10503
10503-097339

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANCCCC
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTcantt
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTcagagatg
TTTAATGCATATTTAANTTATTTAATGTATTTcATNTcATGTTTTCTTATTGTcACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

09789-10150

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTTCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

097559-0100
"0510" 65232560

FIGURE 117

GCAGCTCCGGGTGCTGTGGCCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCCCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAACAAGCCTGGCAGGGTCTCACTTTGTTGCCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAAATGGAAAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGGACCCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCAGTTGAGAAA
GGTACCCCGGATTGTGAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTTCAGCATCTTGACAAA
AGTTTCTTAACCAATTTCCCTTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCTCAGCATGTTCTAACTGCTGCCACTGTGTTTATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGATCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGGCCGGGGTTC
AGAGGATTGCCGAAGGGAGGCCTTCTTTCAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACCTGGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGTCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTTCGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAAGTGGGTGGATGTCCACGGGGTTCAGAAGGAC
TACAACGTTGCTGTTTCGCATCACTCCCTAAATACGCCCAGATTTGCCTCTGGATTACGG
GAACGATGCCAATTGTGCTTACGGCTAAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCAACATTTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACCTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTGAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTTAGGACAAAGAATCTGTAAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCTTAAAAATGAATAAAATTTATGAATATGA

[illegible]

<subunit 1 of 1, 413 aa, 1 stop

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFAEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFEENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNPFFSTAVKLSTGCSGILISPQHVLTAACHCVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKSGRGRQRIAEGRPS
FQWTRVKNTHIPKGWARGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGM IHFS
GFDNDRADQLVYRFCVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HOWVDVHGVOKDYNVAVRITPLKYAOICLWIHGNDANCAYG

Signal peptide:

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

amino acids 236-239

Serine proteases, trypsin family, histidine active site:

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAAGGTCCTGATCACTT
ATATCTGGAAACCAAACCCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTTCGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATAACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACTTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCTCAGAGGAGCCCTCGTAAGTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAAAAAA

[illegible]

```
<subunit 1 of 1, 525 aa, 1 stop
```

MECCRRATPGTLLLLFLAFLLLSSRTARSEEDRDGLWDAGWPWSECSRTC GGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCO
AKGTTLVVELAPKVLGDTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETCTLQGTKGENSLSSSTGTFL
VDNSSVDFQKFDPKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPECTVTGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQELEEGA AVSEEPS

Signal peptide:

N-glycosylation site.

Thrombospondin 1

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTCTGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGGCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGACTCCTCAAAAGTTTCAGTCCCAGCA
GGATTGTGGTAGTTTCTTCCAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAACCTGGCTAACATTCTTTT
TACCAGGGAAGTAGCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTTCATGGGCTTTTTTCAAACTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAATAGGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTTGAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGGAATAAAATTTACTGGTAC

1.000	
2.000	0.999
3.000	0.998
4.000	0.997
5.000	0.996
6.000	0.995
7.000	0.994
8.000	0.993
9.000	0.992
10.000	0.991
11.000	0.990
12.000	0.989
13.000	0.988
14.000	0.987
15.000	0.986
16.000	0.985
17.000	0.984
18.000	0.983
19.000	0.982
20.000	0.981
21.000	0.980
22.000	0.979
23.000	0.978
24.000	0.977
25.000	0.976
26.000	0.975
27.000	0.974
28.000	0.973
29.000	0.972
30.000	0.971
31.000	0.970
32.000	0.969
33.000	0.968
34.000	0.967
35.000	0.966
36.000	0.965
37.000	0.964
38.000	0.963
39.000	0.962
40.000	0.961
41.000	0.960
42.000	0.959
43.000	0.958
44.000	0.957
45.000	0.956
46.000	0.955
47.000	0.954
48.000	0.953
49.000	0.952
50.000	0.951
51.000	0.950
52.000	0.949
53.000	0.948
54.000	0.947
55.000	0.946
56.000	0.945
57.000	0.944
58.000	0.943
59.000	0.942
60.000	0.941
61.000	0.940
62.000	0.939
63.000	0.938
64.000	0.937
65.000	0.936
66.000	0.935
67.000	0.934
68.000	0.933
69.000	0.932
70.000	0.931
71.000	0.930
72.000	0.929
73.000	0.928
74.000	0.927
75.000	0.926
76.000	0.925
77.000	0.924
78.000	0.923
79.000	0.922
80.000	0.921
81.000	0.920
82.000	0.919
83.000	0.918
84.000	0.917
85.000	0.916
86.000	0.915
87.000	0.914
88.000	0.913
89.000	0.912
90.000	0.911
91.000	0.910
92.000	0.909
93.000	0.908
94.000	0.907
95.000	0.906
96.000	0.905
97.000	0.904
98.000	0.903
99.000	0.902
100.000	0.901

<subunit 1 of 1, 336 aa, 1 stop

MAVATAAAVLAAALGGALWLAARRFVGVPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEAAAGQLRREL RQA AECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNH LGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSR SKLANILFTRELARRLEG TNVTNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Signal peptide:

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTGGGATAAGAGAATTTTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGA CTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACCATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACCTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCACTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACCTATAGAGTTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAACCCGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCCTGGAAAACTCAAGCAATTG
TATTTGACTTACATTCTGGAAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFPTVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKEYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLI PGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

Downloaded from www.sagepub.com at 09:23 09 June 2015

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

0978259101201

FIGURE 128

GCCCACGCGTCCGATGGCGTTCACGTTGCGGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTACTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTGTCATGTTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT
TGTTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTTGCTTGTGGAAAGACTG
TTTTCATATGTTATACTCAGATAAAGATTTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTGACAGGTTTGAACCTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTTTATAGGAACTTGTA
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCCCTTTAGT
GCAATACAATAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 131

CGGACGCGTG GGGGAAACCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGCGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCACCTCCTCATAGAGCTTTTAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSLWVRTQLGLPPLLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQCLTYPLHTYP
KEEELYACQRCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQQLMSLMPKMHLFPLTLVRSFWSMMDSAQSFITSSWTFYQLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTGCG
TGAAGTGAAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

09575259 " 10.1501

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGCCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGT CAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGGCCGCGGCTG
GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGAAGAAACTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTG
CATTTTGATGAGAATTCATTTTTTTGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACT
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAARCFQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNP
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

00976250 "04504
"00070" 6528650

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTACCATT
CTTNGAGCGCCCAGATTTTCAACTNNTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACCTCCCCTTCGTCACTCACCTGTTCTTGCCCCCTGGTGTTCTTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCTCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTCGCGTG
AGCCAAGATTGTGCCTCTGCACCTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCTTCCAACTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACCTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCTTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAAC TGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGT TAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AAATGCAGACTTTTCACAATGGTCTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATAACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CAATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCTGGTACTGGCCCTGTTTGCCCTTTGTTGGCTTCATGTGATCCTTGT
GGTCGTGCCACTGTTCTGTCTGGAATAAGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCAGACACCTTGAAAATAACCAATTACCCAGAGTTAATCAGTGCAGGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACCTCCTCAGGCCCTGGAT
CTCATAGGTTTGCAGGAAGGGCCAGGTGAAGCCGAGAACCCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGAGAGAGCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACCTTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLWMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMLILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCTT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCTCCTTTACAGCCACCGACTCAGTGAAGTGCGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAACTCTACGGGAATTTACAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAAACAGCCGGCCCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTTGAGACTCCCCTTCGAGGATTGCACCCGCCCCGTCTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLFLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLNLRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC
 GTTCGTGTCCCCGCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAA
 GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTGAGGGAGC
 CGCAGAGGCGGAGGCTCGCGTATTCTGCAGTGAGCACCCACGTCGCCCGGACGCTCGGGTGTCTCAGGCCCTTC
 GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTGGCT
 CACCTCTCCCAGGAACTTCACTGGAGAGCCAAAAGGAGTGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
 ATCCTGAGGTCATTCATTATGAAGTGTACCGCGCGGGAGTGGCTCAGAGTAACCACAGTGCTGTTTCATGGCTAGA
 GCAATTCAGCCATGGTGGTCCCAATGCCACTTTATTGGAGAACTTTTGAAAAATACATGGATGAGGATGGT
 GAGTGGTGGATAGCCAAACAACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
 AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
 TCTGCAGAATCCTGGGCTGAAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
 TTGGGAGCACACTGGGGAAGATATAGGCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
 AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTCAGGTGTTCTGGCCCTGTATGTACACATTATACA
 CAGGTGCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTACATAACATGAACATCTGGGGGACG
 ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAACTGGTGGGGCCATGCCCTTACAAA
 CATGGGCGGGCCCTGTTCTGCTTGCCACCTAGTTTTGGAGGGGGCTGTAGAGAAAATCTGTGTACAAAGAAGGG
 TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
 CATGTCGGACAAGATCAGATGATAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCAAATTGTTTTCT
 TGTGAAGTAAGATTAAGAGATCAGTGCAAGGAACAACCTGCAATAGGTACGAATGTCTGCTGGCTGTTTGGAT
 AGTAAAGCTAAAGTTATTGGCAGTGATACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTATTATGGT
 ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCACTAAGTCCAATAGA
 AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACAGTTTCAGGCTGTG
 ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCTC
 CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGAATTGGAACTCGAGTTTATTCTGATCTGTCCAGTATC
 TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
 AAGACCTACATTGCTTCTTTTCAAGATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT
 AGAGTGTTTTGCTGTGTGTGAAGTGAATACTTGAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
 ATTTTGTATAAACTGTAACTACTGTACAGAGTACATCACTATTTTTCAGCCCAAAAAGGTGCCAAATGCATA
 TAAATCTTGATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG
 TTTTAGAAATCCTGTGTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
 GTTCTACGTTTCATATATTATATGGTGTCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
 GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAACTGAAAGAAACCTTATCACATTTTCCCC
 AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC
 TGTTAATTTAGGCATATAGAATATTAAATCTGATATTGCATTTCTATTTTATATAAAATAATCCTTTAATATC
 CAAATGAATCTGTAAAAATGTTTGATTCTTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
 ATGAAAACATTCCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
 AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
 GAACAAAGATGAACTAATGTATTACATTACCATGGCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA
 ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
 AATCTGTAAAAATGTAGTTTGGTAATTTTTTTCTGCTGGTGGATTACATATTAAATTTTTTCTGCTGGTGGAA
 TAAACATTAAAAATTAATCATGTTTCAAAAAAAAAAAAAA

[illegible]

<subunit 1 of 1, 500 aa, 1 stop

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLLEKYMDEDEGEWWIAKQRGKRAITDNDM
 QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
 GRYPPTTFHVQSWYDEVKDFSYPYEHECNPYCPCFRCSGPVCTHYTQV VWATSNRIGCAINLC
 HNMNIWGQIWPKAVYLCNYS PKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRY
 YPPREEETNEIERQQSQVHDTHVRTRSD DSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
 YEC PAGCLDSKAKVIGSVHYEMQSSI CRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI
 QTIGKYQSANSFTVSKVTVQAVT CETTVEQLCPFHKPASHCPRVYCP RNCMQANPHYARVIG
 TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSES LQNPPGGKA FRV
 FAVV

Signal peptide:

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCG
CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTGAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTTGCTCCATTCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT
GGTTAATTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGTTTTCATATTCTTTCTAAGATGGAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

0092229-101301

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTACCCCGGTCTGCGTCA**ATG**TAAACTCCAATGTCCTCCTGTG
 GTTAACTGCTCTTGCCATCAAGTTCACCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
 CAAAATCCGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCTA
 TGCCCTACCCCCACTGGAGAGAGGGCGGTTTTCAGCCCCAGAACCCCCGTCTCTGGAAGTGGCATCCGAAATAC
 TACTCAGTTTGCTGCTGTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
 TACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTAAACATCTACGT
 GCCCACGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
 TATTCATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAA
 CATGATTGACGGCAGCATTTTGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACT
 AGGGTTTTTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
 GATTGAGGAGAATGTGGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
 CTGTGTCAGCCTGTTGACCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTAGAGCGGCACCGC
 CCTGTCCAGCTGGGCAGTGAACCTACCAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTGGTGGTGAACAT
 GCTGGACACCAAGGACATGGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
 GGCCACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGACGTCTCCAGACGACCCCCAGATCCTGATGGA
 GCAAGGCGAGTTCCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCTGAAGTTCGTGGACGGCAT
 CGTGGATAACGAGGACGGTGTGACGCCAACGACTTTGACTTCTCCGTGTCCAACCTTCGTGGACAACCTTTACGG
 CTACCTTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAACCC
 GGAGACGCGGCGGAAAACCTTGGTGGCTCTCTTTACTGACCAACAGTGGGTGGCCCCGCGGTGGCCGCGACCT
 GCACGCGCAGTACGGCTCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAGCGAAATGAAGCCCAGCTG
 GGCAGATTCGGCCCATGGTGATGAGGTCCCCATGTCTTCGGCATCCCATGATCGGTCCACCGAGCTCTTCAG
 TTGTAAC'TTTTCCAAGAACGACGTATGCTCAGCGCCGTGGTTCATGACCTACTGGACGAACCTTCGCCAAACTGG
 TGATCCAAATCAACCAGTTCTCAGGATACCAAGTTCATTACACAAAACCCAACCGCTTTGAAGAAGTGGCCTG
 GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
 AACGAAAGTGGCTTTCTGGTTGGAACCTCGTTCCTCATTTGCACAACCTTGAACGAGATATTCCAGTATGTTTCAAC
 AACCACAAAGGTTCTCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC
 AACCACCAAACGCCAGCAATCACTCCTGCCAACAATCCCAAACACTCTAAGGACCCTCACAAAACAGGGCCTGA
 GGACACAACCTGTCTCATTGAAACCAAACGAGATTATTCACCGAATTAAGTGTACCATTTGCCGTGGGGCGTC
 GCTCCTCTTCTCAACATCTTAGCTTTTGCGGCGCTGTACTACAAAAGGACAAGAGGCGCCATGAGACTCACAG
 GCGCCCCAGTCCCCAGAGAAAACACCACAAATGATATCGTCCATCCAGAACGAAGAGATCATGTCTCTGCAGAT
 GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGA
 CTACACCTCACGCTGCGCCGGTCGCCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAA
 CACACTGACGGGATGCAGCCTTTGCACACTTTTAAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
 CGGACATTCACCACTAGAGTAT**AG**CTTTGCCCCTATTTCCCTTCCCTATCCCTCTGCCCTACCCGCTCAGCAACAT
 AGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTTGTCCCACT
 GACTTAAGACAAAAATGCAAAAAGGCAGTCATCCATCCCGGCAGACCCCTTATCGTTGGTGT'TTTCCAGTATTAC
 AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGT'TAAAATAACTGCTTTAAGATCTCTACCA
 CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAGGCCCCGGGTGTTTCCAACGTCATGGAAGCAGCT
 GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTAAATTACAATGGAAGTTTAAACATTTCTTTCTGTGC
 CACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGCCAGCACATGGAGCTGTAATCCAG
 AGAGAAGGAAACGTAGAAAATTTATTATTAAGAAGTGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAGAG
 GTGTTTTTGCCAGCCTGAACATATATTTAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVYPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCMNL
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYPDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKP SWADSAHGDEVYPYV
FGIPMIGPTELFSCNFSKNDVMLS AVVMTYWTNF AKTGDPNQPV PQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPDMS
FPYGTRRSPAKIWPTTKRPAITPANNPKH SKDPHKTGPEDTTVLIETKRDYSTE LSVTI AVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTTCGAGTACTTGAAACGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCCTACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCTGAGAGACTGGGAGTTGCAGGTGCACCTTCAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCTGTTGATTCTGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAAGTACAGTGGAGAGAACCCAGAAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCATGCCGCC
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGTGTCTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCATG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACTGTGTTGTGATGCCAAATCACAGAAC
AGAATTTTCATAGCCAGGCTGCCGTGTTGTTGACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCTGTTGGGTT
TTGCATTTGACCCAACCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAACT
TCTTCCCTGCCCTTACCTTCTTTTCACTCCATTCAATTGTCCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCTGGGGCTGCAGAACACACCTGCGTTTTAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAAGTCTGAGTCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGQPQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFGVGLGVFVDITYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEFTVERTPEEEKLHRDVFLPSVDNMKLPENTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCTGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAAATCAGTTGTCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAAACCATGAAGAGAAAAATAGAATACCTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCATGCC
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTTATTGCGCACGGAGACAGGTACCCACTGTATGTATTCCCA
AAACAAAGCGACCAGAAAATTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAAC TGGAAGCTTTCA
TTAGTCACATGTCAAAAGGATCCGGAGCCTCTTTGAAAGCCCCCTTGAACTCCTTGCCCTCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCGACATTTGCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAACTCCTGCCCAATGATTGGTCTGCAGAC CAGCTCTATTTAGAGACCACTGGGAAAA
GCCGGACCCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCAGATTTTGACTGGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGGCTGCTGCTGCTGGAAGCTGCTATTGCCCGGTAAGAAAACAGTATCTGGAAGGAGC
AGCGTCGTGAGTACCTCTACGTTTGAAAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTCAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAAGAAATGTACTTCGGGTATTCTCTCTGGGTGCCACCCCATCCTGAACCAACCATCG
GCCGGATGACGCTGCCACCGAGGGCAGGAAAGAAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCTTGGGCCTTTGAGAAGCCAGGTTCCCAAGGTTTGAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCAGTGAACATTCCGTCCGGATTCTTTACAATGGCGTCGATGTCAATTCCACACCT
CTTTCTGCCAAGACCACCAAGCGTTCTCCCAAGCCCCATGTGCCCGCTTGAAAACTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTAATATGATGCATGTCAAGGGAAGGATTC TAAAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC
TAAGGGTAGAAGATTATTGCTTTTTTAAAGGCTAAATATTGTTTGTGGGAACCACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC
TCTTCTGGCCTGCCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAAAATTTTTTTAATCTTAGACATT
TTTACCTTGCTCCTTGTTAAGAATTTCTTGAAGTGATTTATCTAAAAATAAAGGTGGCAAACTTTTCTGTAAAGG
GCCAGATTGTAAATATTTTCAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTCTATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTCCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCTGGGCTGTAGTTTGTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAAATTAATCAATGGTTTCTCCAGTAACCTCTGCTAGAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTTTATGATGATTGTGTGTAACAAGATAGTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAATATT
TTCTAAGACCAGTTTATAGTACTCTTATTCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACTTTCTTCCAATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA
ACTGGATTCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTTGATTGATTTTTTAAATGCGTTTTTGGGA
AGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACCTAATGGTGCTAATTCAGAGAAATGGAAAGTGAAAGTGAGATTCTCTGTTGTCTCGGCATTCC
AACTTTTTCTCTTTGTTTTTGTCCAGTGTTCATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAAGATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNI PSVAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC
TRNGCVDMEHFVKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLNLVRFVKRDMFVALGGSGTNYYDACHREGF

Signal sequence:

amino acids 1-18

09973299-101501
"T05T0" 6294660

```

1  # 2019-08-26 10:00:00
2  # 2019-08-26 10:00:00
3  # 2019-08-26 10:00:00
4  # 2019-08-26 10:00:00
5  # 2019-08-26 10:00:00
6  # 2019-08-26 10:00:00
7  # 2019-08-26 10:00:00
8  # 2019-08-26 10:00:00
9  # 2019-08-26 10:00:00
10 # 2019-08-26 10:00:00
11 # 2019-08-26 10:00:00
12 # 2019-08-26 10:00:00
13 # 2019-08-26 10:00:00
14 # 2019-08-26 10:00:00
15 # 2019-08-26 10:00:00
16 # 2019-08-26 10:00:00
17 # 2019-08-26 10:00:00
18 # 2019-08-26 10:00:00
19 # 2019-08-26 10:00:00
20 # 2019-08-26 10:00:00
21 # 2019-08-26 10:00:00
22 # 2019-08-26 10:00:00
23 # 2019-08-26 10:00:00
24 # 2019-08-26 10:00:00
25 # 2019-08-26 10:00:00
26 # 2019-08-26 10:00:00
27 # 2019-08-26 10:00:00
28 # 2019-08-26 10:00:00
29 # 2019-08-26 10:00:00
30 # 2019-08-26 10:00:00
31 # 2019-08-26 10:00:00
32 # 2019-08-26 10:00:00
33 # 2019-08-26 10:00:00
34 # 2019-08-26 10:00:00
35 # 2019-08-26 10:00:00
36 # 2019-08-26 10:00:00
37 # 2019-08-26 10:00:00
38 # 2019-08-26 10:00:00
39 # 2019-08-26 10:00:00
40 # 2019-08-26 10:00:00
41 # 2019-08-26 10:00:00
42 # 2019-08-26 10:00:00
43 # 2019-08-26 10:00:00
44 # 2019-08-26 10:00:00
45 # 2019-08-26 10:00:00
46 # 2019-08-26 10:00:00
47 # 2019-08-26 10:00:00
48 # 2019-08-26 10:00:00
49 # 2019-08-26 10:00:00
50 # 2019-08-26 10:00:00
51 # 2019-08-26 10:00:00
52 # 2019-08-26 10:00:00
53 # 2019-08-26 10:00:00
54 # 2019-08-26 10:00:00
55 # 2019-08-26 10:00:00
56 # 2019-08-26 10:00:00
57 # 2019-08-26 10:00:00
58 # 2019-08-26 10:00:00
59 # 2019-08-26 10:00:00
60 # 2019-08-26 10:00:00
61 # 2019-08-26 10:00:00
62 # 2019-08-26 10:00:00
63 # 2019-08-26 10:00:00
64 # 2019-08-26 10:00:00
65 # 2019-08-26 10:00:00
66 # 2019-08-26 10:00:00
67 # 2019-08-26 10:00:00
68 # 2019-08-26 10:00:00
69 # 2019-08-26 10:00:00
70 # 2019-08-26 10:00:00
71 # 2019-08-26 10:00:00
72 # 2019-08-26 10:00:00
73 # 2019-08-26 10:00:00
74 # 2019-08-26 10:00:00
75 # 2019-08-26 10:00:00
76 # 2019-08-26 10:00:00
77 # 2019-08-26 10:00:00
78 # 2019-08-26 10:00:00
79 # 2019-08-26 10:00:00
80 # 2019-08-26 10:00:00
81 # 2019-08-26 10:00:00
82 # 2019-08-26 10:00:00
83 # 2019-08-26 10:00:00
84 # 2019-08-26 10:00:00
85 # 2019-08-26 10:00:00
86 # 2019-08-26 10:00:00
87 # 2019-08-26 10:00:00
88 # 2019-08-26 10:00:00
89 # 2019-08-26 10:00:00
90 # 2019-08-26 10:00:00
91 # 2019-08-26 10:00:00
92 # 2019-08-26 10:00:00
93 # 2019-08-26 10:00:00
94 # 2019-08-26 10:00:00
95 # 2019-08-26 10:00:00
96 # 2019-08-26 10:00:00
97 # 2019-08-26 10:00:00
98 # 2019-08-26 10:00:00
99 # 2019-08-26 10:00:00
100 # 2019-08-26 10:00:00

```

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCATCCCCTTTTGAAGAACAGTACTGTGGA
 GCTATTTTAAGAGATAAAAAACGAAATATCCTTTCTGGGAGTTCAGATTGTGCGAGTAATTGGTTAGGACTCTGAGC
 GCCGCTGTTTCACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGACGCGCCTGAAGCACAAAGCAGAT
 AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCATTCTAT
 GGGCGAAGGAAC TGCTCCTGACTTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAATGAT
 TCCTGCGCGACTGACCCGGGACTACAAGGGCTTGCTCTGCTGGGAATCCTCTCTGGGACTCTGTGGGAGACCGG
 ATGCACCCAGATACGCTATTTCAGTTTCGGAAGAGCTGGAGAAAAGCTCTAGGGTGGGCGACATCTCCAGGGACCT
 GGGGCTGGAGCCCCGGGAGCTCGCGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCGAGCTTTTCGCCC
 GAATCCGCGCAGCGGCAGCTTGGT CACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
 TCAATTAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
 CGACAATGCGCCCTTACTTTCTGTGAAAGTGAATTAGAAATAAAAAATTAGTGAAAAATGCAGCCACTGAGATGCGGTT
 CCCTCTACCCACGCTTGGGATCCGGATATCGGGAAGAACTCTCTGACAGCTACGAGCTCAGCCGGAACACTCA
 CTTCTCCCTCATCTGCGCAAAATGGAGCCGACGGTAGTAAGTACCCCGAATTGTTGCTGAAACGCGCCCTGGACCG
 CGAAGAAAAGGCTGCTCACCACTGGTCTTACGGCCTCCGACGGGGGCGACCCGGTGCACAGGCACCGCGCG
 CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCAGCGTTTGCTCAGCCCGAGTACCGCGCGAGCGTTCC
 GGGAATCTGGCCTTGGGCACGCGAGCTGCTTGTAGTCAACGCTACCGACCCCTGACGCAAGGAGTCAAGTGCGGAATG
 GAGGTATTCTTCCGTTATGTGGACGACAAGGCGGCCCAAGTTTCAAATAGATTGTAATTCAAGGACCAATG
 AACATAGGGGAGTTGGACACAGGAGTCAAGATTCTACAGATGGAAAGTGCAGCAATGGATAATGCAGGATA
 TTTGCGCGAGGCCAAAAGTCTTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCTCACCTCTCT
 CGCCAGCTCGGTTCCCGAAAACTCTCCAGAGGGACATTAAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA
 GGAAAACGAGCAGGTGATCTGTTTCATCCAAGGAAATCTGCCCTTTTAAATTAGAAAAATCTTACGGAAATTACTA
 TAGTTTGTAGTCACAGACATAGTCTTGGATAGGGAACAGGTTCTTAGCTACAACATACAGATGACCGCCAGTACGG
 GGAATCCCGCCCTTATCCACGGAATCATATCTCGTGAACGTGGCAGACCAACGACACCCCGCGGTCTTT
 CCCTCAGGCTCTCTATTCCGCTTATATCCACAGAGAACAAATCCGAGGAGGTTTCCCTCGTCTCTGTGACCGCCA
 CGACCCCGACTGTGAAGAGAACGCCAGATCACTTATTTCCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC
 GTCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGTCTCTTCGACTACGAGCAGTTCCGAGA
 CTTGCAAGTGAAAGTGATGGCGCGGACAAACGGCTACCCGCCCTTCAGCAGCAACGTTGTCGTTGAGCCTGTTCTGT
 CTGGACCAAGAACGACAATGCGCCGAGATCTGTATACCCCGCCCTCCCAAGGACGGTTCCACTGGCGTGGAGCT
 GGTCCCCGCTCCGACAGGCCCGGCTACTGGTGACCAAGTGGTGGCGGTGGACAGAGACTCCGGCCAGAAGCG
 CTGGCTGTCTTACCCTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
 CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCGCTCCAGGACCACGGCCAGCC
 CCTCTCTCCGCCACTGTCAAGCTACCGCTGGCCGTGGCCGACAGACTCCCCCAAGTCTCTGGGCGACCTCGGCAG
 CCTCGAGTCTCCAGTCAACTCTGAAACCTCAGACCTCACTCTGTACTGGTGGTGGCGGTGGCGCGGCTCTCTGT
 CGTCTTCTGGCCTTCGTCACTTGTCTGTGGCGCTCAGGCTGCGGCGTGGGACAAGTCAAGCCTGTGTCAGGC
 TTCAGGAGGCGGCTTGACAGGAGCGCGCGCGTGCACCTTTGTGGGCGTGACAGGGGTGCAGGCTTTTCTGCAGAC
 CTATTTCCACGAGGTTTCCCTCACCAAGGACTCGCGGAAGAGTCACTGATCTTCCCCAGCCCAACTATGCAGA
 CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAGCGAGCCCTTTTGTCTGTCAGGTGATTGGGATTTTCTAAAGA
 CAGTCATGGGTTAAATTGAGGTGAGTTTATATCAAACTCTTCTTTTCTTTTAAATTGCTCTGTCTCCGAAGC
 TGGAGTGCACGGGTACGATCATAGCTCACTGCGGCCCTCAAACTCTTAGGCTCAAGCAATTATCCACCTTTGGCT
 CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCATCTATCTATCTATCTATCTATCTATCTAT
 CTATCTATCTATCTATCTATTACTTTCTGTACAGACGGGAGTCTCACGCCGTGAATCCAGTACTTTGGGAGGC
 CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGGAGACCAGCTGACCAACATGGAGAAACCCCGTCTATACTAA
 AAAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCAGCTACTTTGGGAGGCTGAGTCAGGAGAAT
 TGCTTTAACTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG
 AACCTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVP EELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFP LPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELV LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLIITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYISLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSF DYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILL LALRLRRWHKSRL LQASGGGLTGAPASHFVGVD
GVQAFLQTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQT PRLKQLSHLCLRCNRD YRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGSLTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCAGAGCAAGAAATACCTGTGGTTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCCTCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACCTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCCG
CGCGCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCGCGCCCCG
GCCCCGCGCCAGGTGAGCGCTCCGCCCCGCGGAGGCCCGCCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACCTGATCCCATAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGGGCCGTAGCGGCGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTTCAGTCCCCGACGCGC
CCCGCCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGC
GGTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGACCCTGCCACATTCACGGCCTGGGCGCCCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCCTGCACGGCAACCGCATCTCCAGCGTGGCCGAGCGCGCCTTCCGTGGGTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCCTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCCCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGGTGGCGGGACTGGTGAATCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGAACCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCCTCCCTGATGGACGCCTGCCGCCCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCAGAACGCCGCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCLQELGPGFLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRRLNDNPWVCD CRARPLWAWLQKFRGSSSEVP CSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCTCTGGAGAGGACTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAACACAGCCTCAGATACTGGGGACTTTAC
AGTCCCACAGAACCGTCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCCTTTTCTCTCTCTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAACTGAGGGCAGCTCCTTTGTAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGTTAGGGTTG
TTTCCAGAGGGAACAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTACACAGAGCCCTGTGTCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTCTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCCTCCTGGG
ACTACGTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAATCCTATTTTTCGGGTCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGGAGGAAAGCTGAGCTCAGGTTAACA
CTCACAGCACTGGATGGTGCTCTCCGCCAGATCTGGCACTGCTCAGGTTACATCGAAGT
CCTGGATGTCAACGATAATGCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAATCTCTG
AGGACAGTCCGGTAGGCTTCTGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTTCTATTCATTTTCCAAGCTTCAGAAGAGATTGGCAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACATAAAAAACAACCTCGATTTGAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGAAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACCGCCTGAACTGTGGTTGCACTTTTCAGTGTTTCAGATCTTGATTACAGGAG
AAAATGGGAAAATTAGTTGCTCCATTCAGGAGGATCTACCCTTCTCCTGAAATCCGCGGAA
AACTTTTACACCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTACCCAAACCTCCTACACCTGTTTCGTC
CGCGAGAACAACAGCCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCCAGGTACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTTCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTACAGACCACGGCTCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTGCTGGACGCCAACGACAACCTCGCCCTTCGTGCTGTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCGGGCGGCCGAGCCGGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTCGTACCA
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTTCGGCGTGTGGGCGCACAAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGCTGGTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCGACCCAGGCCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCTCTTTTCGGTGCTCCTGTTT
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCTCGGTGGGTGCTGCTTGGTGCCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCTATCCAGAGCTACC
AGTATGAGGTGTGCTGTCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
ATCCCCAATTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
TAACTTTGGGTTCAATATTCACTGAACCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
TGTGGCATTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTCTCATGTTCTTCTCCCTTTGTTTTAAAGTGAACATTTACCTTTATT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRQVLLSPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFVLVVKVSATDVDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGFLGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAA
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMMSGTRT
LSQSYQYEVCLAGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCCGGTGCCGGTGCACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GCCGCTTTTCTGCCAGTTTGGCCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCCCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATAAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCTTATTCTTTTTTAATTAGTTTTATTATGTATGCTACCATTCGAA
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFTYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETCLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFLLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTGCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGGGCTCTTCCTC
TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAAGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCTGGCCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCCTGAAAAATAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCC GG CAT CCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAAC TAATA
AAATCATGAATATTTTAA

[illegible]

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
```

```
><subunit 1 of 1, 295 aa, 1 stop
```

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLFLASHCCLGSARGFLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLVGVSERDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGELVITSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTGCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTCGAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACC
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCTGA
GTGTTCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTGTAAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

097629-101501

[illegible]

><subunit 1 of 1, 560 aa, 1 stop

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVGTON
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGLVLVANVLSPLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQIQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAAATGMIFVLGQAEGILIMLAMTALTVRREPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Signal peptide:

Transmembrane domains:

N-glycosylation site.

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFVDLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALAIQEDAAPGTTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLTPTLYRTLNRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEPESPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAAAGMKVQGDPPGKGTGTEGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

"TQST" 652460

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTTACATTGGC
TATTTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCATTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT
TCCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCTTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAACTCACTGAAGTCCTTGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTACCTTCATAGCATACTCCTTCCCCGTGAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDVPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

[illegible]

CATGGGAAGTGGAGCCGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCAACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

[illegible][illegible]

FIGURE 176

CTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCTCCGGGATTGCTACCTTTT
TGGCTCCCTGCTCGTGAACCTGCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCGCAA
GGAGGGGCGAGCCAGGCAGCCTCTTCCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCGAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTA CCTGTGCACACCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTCGCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGAATGGAAGTTCTGTGAGGGACGCCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCTTCTCCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGA CGACGG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCAACCA
CAAGGGTGTGTGGTCTCTGCGCAAGGACAGCGCCAGTGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACCAGGGGGGTCACTG
GGCTGGGATCTCCCCCTCTCCGGCTCTGCGGCTCCCCCTGACTCCATGTTCGGGATCAGCCTGGCTGTCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGGTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCA
TGGGAGCAGCCTGGGGGTTGTGCGCAACCTTCAACAGTGTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGTCTTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCTGTAACCTGGAAGAACCACAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGT
AGTGACCTTGTCTTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC
CCCCATCCTCAATGCCACCAGCCAGCCAGCCAGCGGGCAGAGATCCACTTCCCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCACACC
TCTGCCCATGGATGTGGATGGAACAAAGCCCTGTTTGACTGAGTGGGCGAGCCAGTCATTGGCCCTGGAGCTGAT
GGTCACCAACCTGCCATCGGACCCAGCCAGCCCGAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGT
CATGCTTCTGACTCACTGCCTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAAGCCA CTCTGCCTGTCCAA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTACCTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAAGTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC
CCCGCATCTTCTTCTCTGCTGTGGTGTGAGGGGAGAGAGCCATGCACTGTGAGCGGGATGTGGGCAGCAAGT
CAAGTATGAGGTACCGGTTTCCAAACCAAGGCCAGTGCCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGG
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGCTG
TGAGAAGAAGAAAAACATCACCTGGACTGCGCCCGGGGCACGGCCAACCTGTGTGGTGTTCAGCTGCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCGCTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA
TGCCTCCACAGTGATCCAGTGATGGTATACTTGGACCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCCTCCTGGCTGTA CTGGCTGGGCTGCTGGTGTAGCACTGCTGGTGCTGCTCCTGTGGAAGATGGGATTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAAC TGGGGCAGCCCCCGGCGGGAGGGCCCGGATGCACA
CCCCATCCTGGCTGCTGACGGGCATCCCCAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCTTAGGTTCC
CATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTGCATCAAGATTTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAAC
TCTCCACCCAACTTCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGCTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGT
GTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCTTAACCTAGAGGGTTCGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGCTCCTTCTCTAGTTTCCCTCTCATCTGACCTTAGTTTGTGTCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAATAAATATTTGAGAACAATAAAAAAAAAAAAAAAAAAAAA

09975299-101501

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFC
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKG TARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGD LNQDGFPDI AVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFYSLSGSLDMDGNQYPDLLVGS LADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTD RRLRGQVPRVTFLSRNLEEPK
HQASGT VWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQ PSTQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSPDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRA LDPAEKPLCLSN
ENASHVECELG NPMKRG AQVTFYLILSTSGIS IETTELEV ELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVT VSNQGQSLRTLGS AFLNIM
WPHEIANGK WLLYPMQVELEGGQGPQKGLCS PRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANC VVFSCLY SFDRAAVLHVWGR LWNSTFLEEY
SAVKSLEVI VRANITVKSSIKNLMLRDASTVIPVMVYLDPM AVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYH AVKIPREDRQQFKEEKTGTILRNNWGS PRREGP
DAHPI LAADGHP ELGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAACAAGATGCTCAAGGTGTCAGCCGTAAGTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCTTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGACACATACAAGGACAGTTTAAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTTCAGATTGTGCTATAGATTTTGAGATCTCCGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTATGAC
CATGATGTATACATTTGAATTGATGACAGTTGAAATCAATAAATCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTACGAGAACAAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLMKKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSILPI
CKDSLGMFMNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDDEDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAAACAGTACCTGACGC
CTCTTTTCAGCCCCGGGATCGCCCCAGCAGGGATGGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
GCTCTGCCCTCCGGTGCTGCTGCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCCTT
CCCCCGGCCAGAAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTGA
GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
TCAGATGGAGTTTCACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTCAGCACCATT
TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAGATTGGAAG
AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATTTCTGCTTAGAGCATTGTAAGCTCGTGATCGAAACATACAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTTCATGGTGGTGGTGTGAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTTGAAGATAAGAGGAAAAAGTAGAACTTAAAACTCCAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATA
AAAGTAGGAAACAGGTATAAATTTAATGTGAAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
TTTCCTAACTTTGAAAAATTTTGCAAATGTCTTAGGTGATTAAATAAATGAGTATTGGGCCTAATTGCAACACC
AGTCTGTTTTTAACAGGTTCTATTACCCAGAACCTTTTTGTAAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
TCAGTTTTAAGTTATAAATCACCTGAGAAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
CTTTTCTCTATTACATATGCATCTCTCTATAATGTAAATAGAATAATAGCTTTGAAATACAAATAGGTTTTTTG
AGATTTTTATAACCAAATACATTTTCACTGTAACTATTTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTTACATT
CCAAAAGCTGACATTTTACGATTCTTAAAAACACAAAGTTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
AAATGAAGAATATAGTTTAAAGCTTCTCTCTCATAGGGACACATTTTCTCTAACCTTAACTAAAGTGTAGGA
TTTTAAAATTAATGTGAGGTAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
TAATCATGTTATGTTAATTTAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA
TTGCTAAAATGATCTGGGCCCTACCATAAATAAATATCTCTTTTCTGAGCTCTAAGAATTATCAGAAAAACAGGAA
AGAATTTAGAAAAAAGTTGAGAAAACTAATCCAAAATAAAATTCACTTAAGTAGAACTATAAATAAATATCTAGA
ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AACTTTGGCTGTAGGTTTTATTTTCTACAAGAATCTGGTTTGAATTATTTTGTAAAGCAGGTACATTTTATA
AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAATAAATTTTATTTATAACAACTTTTAT
TAAATGGCCTTTCTGAACACTTTATTTATGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTAA
CACCTAAATGTGAATAACCCATATATACAACAAGTTTCTGCCATCTAGCTTTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATAAATTTTAAAGTTATGCCCATTTTATAACGTTGTTTAT
GACTACATTGTGAGTTAGAAACAACTTAAAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
CTTGATGAGCAATAATGATAACCAGAGAGTGATTTTCAATTTACACTCATAGTAGTATAAAAAAGAGATACATTTCCC
TCTTAGGCCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAATGCCGTAT
ATGATCAATTACCTTAATTTGGCCAAGAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
AGGTCAATAAGATCCTTGCCATGAATACCCCTCCCTTTTGCCTGTGTAATTTGCAATGAGAAGCAAATTTACA
GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAAACTGTGATTAAAGAAATCTA
CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT
ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACCTTCTTACCATATAAAAAACGATAATTGCTT
TATTTGGAAAAAGAAATTTAGGAATACTAAGGACAATTTATTTTATAGACAAAGTAAAAAGACAGATAATTTAAGAGG
CATAACCAAAAAAGCAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
CATATGCTTTTTTAAATTTCACTATTTCAATTTCTAAATTAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTT
AACAGCTCATTTTGTCTTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTT
CATAATGTAGCAGTTACCGTGTTTACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
GTTATGCTGTTCTTTCATGTGAATGTCAAGACATGGAGGGTGTGTAATTTTATGGTAAATTAATCCTTCTTA
CACATAATGGTGTCTTAAATTTGACAAAAATGAGCACTTACAATTTGATGTCTCCTCAATGAAGATTCTTTAT
GTGAAATTTTAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTGTG
CTCAAACTGCTTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATA
AAATTTATCAAAGGAAAA

09973299 301501

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTTFSTISEKVIFTEL
ILDNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

"6523456"

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCAGTGTGTCTCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACTCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTCTTTTTCTTACCTTCATTTCAAGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

6978299-101501

52

[illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible][illegible]

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCCTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAAAAA

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC

```

#-----#
# 1. Importing Libraries -----#
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')

# 2. Loading the Dataset -----#
df = pd.read_csv('data.csv')

# 3. Data Overview -----#
df.head()
df.info()
df.describe()

# 4. Data Cleaning -----#
df.isnull().sum()
df.dropna(inplace=True)

# 5. Data Distribution -----#
df['Age'].hist()
df['Salary'].hist()

# 6. Feature Engineering -----#
df['Age Group'] = pd.cut(df['Age'], bins=[18, 25, 35, 45, 55, 65, 75], labels=['18-25', '26-35', '36-45', '46-55', '56-65', '66-75'])

# 7. Model Training -----#
X = df[['Age', 'Experience', 'Education', 'Age Group']]
y = df['Salary']

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(X_train, y_train)

# 8. Model Evaluation -----#
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r_squared = r2_score(y_test, y_pred)

print(f'MSE: {mse}, R-squared: {r_squared}')

```

<subunit 1 of 1, 125 aa, 1 stop

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP

Important features:

amino acids 1-17

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTTAAACCAGCTATCCAAATGCAGTGAACCTTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTTCTTTATGGAACCTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACCTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWISKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCTGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

0.9975259 "0.9975259"

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATCCCGTGCGCCGCGG
CTGGGCCCGTCCGAGAGTGCCTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTACAGATTAAATATTTTTGGGGACAGATTTGTGATGCTTGATTACCCCT
TGAAGTAATGTAGACAGAAGTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTTAC
TTAAATCAGAACTTGCATAAGAAAGAGAAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTACAGTTTACTTGG
AGTGTCCAAAACTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAAATTTGATGC
TGCTGTTAATTTCTGGAGAACTGTGGTTTTGTAATTTTTTACTCCCCAGGCTGTTTCACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTTGCTAAAGAAGTGGATGGGTACTTTCGAATTGGAGCTGTTAACTGTGGTATGATAGATA
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTTCATTTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATACAACTGCTTTTGTCTGGTATTGGCTGGCTGATCAGTTTTTGTTCAAA
AGGAGGAGATTGTTTGAATTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCAGATTTTGAATCTTTCGGCAACACACTAGAGGATCGTTTT
GGCTCATCATCGGTGGCTGTTATTTTTTCAATTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAAT
AAAAACTCTACTTTAAAAATGATCATATTCAGTTGGCAGGTTTGACTGTTTCTCTGCACCAGACATCTGTAGTAA
TCTGTATGTTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGAAATTCATCATGGAAA
GAAGATTCTATATGATATACTTGCTTTTGCACAAAGAAAGTGTGAATTTCTCATGTTACCACGCTTGACCTCAAAA
TTTTCTCTGCCAATGACAAAGAACCATGGCTTGTGATTTCTTTGCCCTTGGTGTCCACCATGTCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTGACAGTTTATGA
GGGACTCTGTAAATGATATAACATTGAGGCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCACCACCTTCAACGAAC TAGTTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTTCTCCGTG
GTGTCATCCTTGCCAAGTCTTAATGCCAGAATGGAAAAGAATGGCCCGGACATTAAGTGGACTGATCAACGTGGG
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCAGGAAAACGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATTTCCCTGAGAAAT
CTGGGGTCTAGGATTTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTTCACTGAAAAAGTTCTACAAGG
GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTTGAGCT
CTTGCTTAGGATGATTAAAGGAAAAGTGAAGCTGGAAAAGTAGACTGTGAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCCTATCCAACCTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAAATTTCAAGAAGAGCA
GATAAATACCAGAGATGCATCAAAATCGCTGCTTAAAGTGAAGTTGGAATAGGGATGCTTATTTCCCTGAGAAAT
GAGGAATAAGGATGAACCTTTGATAATGTTGAAGATGAAGAAAAAGTTTAAAAGAAATTTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCACTGTACTGCCA
GAATTTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGGTTACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAAACACCTTTAAAAAATATTTAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCACTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAAGAGGTAAC
TAGTTTTTGGTCACTTGTTCTCCTAAAAATGCTATCCCTAACCATATATTTATATTTCTGTTTTAAAAACACCCAT
GATGTGGCACAGTAAACAAACCCGTGTTATGCTGTATTATTATGAGGAGATTCCTTATTGTTTTCTTTCTCTCA
AAGGTTGAAAAAATGCTTTTTAATTTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATTTCTGTAGTTTGTGTATCATCCAGGAAAACCTGAGGGAAAAAATTA
TAGCAATTAAGTGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTTATGATTTTTCTGAAATGCTTTTATAGAAATTTTCCACTGATAGTTGATTTTTGAGGCATCTAATAT
TTACATATTTGCTTCTGAACCTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTTATAGTTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTTTCAAATAGGAAAAATTAATTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACGTAGTTATTCCAGTTACTAGTTTACTGTGAGGGCTGCCTTTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAAATCAAGTATATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTGTTTGA
CTCAAAGAATCACAATTTGTGAGTAAATGATGTTGTTTATGTTTATAATTCAGAGTGTACAGAAATGGTAAAAAT
CCAATCAGTCAAAGAGGTCAATGAATTTAAAGGCTTGCAACTTTTTTCAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKLEDNQGGQYESWNYRYDFGI
YDDDPETITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSPLYVFQP
SLAVFKGQGTKEYEIHGKILYDILAFAKESVNSHVTTLGPNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHS
AEQILEFIEDLMNPSVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPTQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPFEFELLARMIKGKVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

[illegible][illegible]

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFNVTGFTKNPSTRLWVPVLETDEVVRSLLID
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCAGCCCGCCGGGGC
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTA CTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGACACGTCCTTCT
CTAGGCCGCACACGGGGCCCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTGACGAGTTTTCTGGACAAGTTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCA ACTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCTTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTGCGCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTGAAAGCTTTCGCGTTGCTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAAATATTTTCAGGTATTTAATACGA

FIGURE 195

TCGGGCCAGAATT CGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGGCCACAATCCGAGAGGGGCATGCTGGCCCAG
CCACTGGGGCCGCATGGGCCAGCCCGCTGAGGTGCGGGGCTGCGGCAGTGTTCTTGGCCTCCGA
AGCCAACCTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCGAGAGCTGGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCCTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLCDVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
 CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCGCCGAGAAGCCTCGCTCG
 GCGCCCAACATGGCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
 CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGAGCAGAGCCGGG
 TCCAGCCCATGACCGCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
 TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTGAGAATGGGAGGCTTTTGCAAAGAA
 TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
 GCCGCTTCTTTGTCACTCTCTCCAGCATTTTTTTCATGCAAAGGATGGGATATTCCGCCGT
 TATCGTGGCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
 AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
 TTTTGTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTACAGTGACTCTTGGAATT
 CCTGCTTGGTGTCTTATGTGTTTTTGTCTATAGCCACCTTGGTTTTTGGCCTTTTTTATGGG
 TCTGGTCTTGGTGGTAATATCAGAATGTTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
 GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
 GAGGAAAAAGATGATTCAAATGAAGAAAGAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
 GAAAGAAGATCTTGGCGATGAGGATGAAGCAGGAAGAAGAGGAGGAGGACAACCTGGCTG
 CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
 ACCCGGGAGGAAGTAGAGCCTGAGGAGCTGAAGAAGGCATCTCTGAGCAACCCTGCCAGC
 TGACACAGAGGTGGTGGAAAGATCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
 TGTAGATTTAATGATGCGTTTTTCAAGAATACACACAAAACAATATGTCAGCTTCCCTTTGG
 CCTGCAGTTTGTACCAAATCCTTAATTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
 CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
 GACAATCAGGATATAGAAAAACAAACGTAGTGTGGATCTGTTTGGAGACTGGGATGGGAA
 CAAGTTCATTTACTTAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
 AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
 CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGTCATCCTTTTTCTTGTGTAAAGTATTTAT
 TTTTGTCAAATTGCAGGAAACATCAGGCCACACAGTGTCATGAAAAATCTTTCACAGCTAGAA
 ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
 TGCTATGTTTTATTTCTTACCTTTAATTTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
 CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
 TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
 GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
 TTACAAGACAGATTAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTTT
 AAGTTTTCTAAGCAATATTTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
 CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
 TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCTTCAGTGATGTGCTTTTGGTGAAA
 GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTCATCTTCTGTAATC
 TTCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
 GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEDEEEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSKHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGA CTCAGAGATTCTCTTTTGTCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATT CAGGGGCTCTCAAGAA CAATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGTTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATT CAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTT CAGATCAGAACCAAGCTACCCAAGAAAACCCATCTCCAAATTTGTG
TATGGATTACGTGTCACTATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAGAGGAAGGGTGAGAAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACCTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTGACCTACAGGTAGGCTAGTATTTATTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTTATTTTTTTGAGATAGGGTCT
CACTTTGTGTACCCAGGCTGGAGTGACGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCCTCAAGTAGCTGGGACTACAGGTGCAATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACCTCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTTTTTGGAGACAGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTTAGT
AGAGACAGGGTTTCTCCATGTGCGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTTAAATACA
AAAATACATAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAAATATGTTTTATTAGTTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTTCAGTAATTA CTGTAAAATGGTATTATTGGAATGAACTATATTTCTCATG
TGCTGATTTGTCTTATTTTTTTT CATACTTTCCCACTGGTGCTATTTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTTTCAT
TTCAATAAAAAAAA AACTCTTATCAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

05978299-103001

FIGURE 203

GGAAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGAATTCCTCAAGAGTGAATTCCTCGGAGGAAAATGACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCTCTGCTGAGTCTGCTCTTCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG
GAAGACTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCCTTTCCTGCAGGCCACCCCTGCTTCCCGA
TCCTTCCCTGACCCAGGGGCCCTTACCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG
GCTCAGGGCCCCCGCTGTAGCCACTTCTGTACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCAGTGCC
GCCAGCTTACCTTCTCCTTCCACAGTCTCCTCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAACTGACCTCTGTGAGATTATGGGGGACATGGTGTCTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCAGCCACAGCCGGCCTCCAGGACCTGCATCTCACTCCCG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCCTCGAACACTCTTCCAGAGGACGAAAGGC
CGGAGCGGGGAGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCTCGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCCGTG
GTGCTCACTTTCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTGAAGACCCACA
TTGAGCAGCCCGGGGCATTGGAGCAGTGTGGGTGTGAGACCCCTCAGGAGAGAAACCCAAACATCCTGCTTCTGC
AACCATTGACCTACTTTCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTGCTCTCTGCCCTGGCTGCCCTTGTACCATTTGCCGCTACCTCTGCTCCAGG
GTGCCCCGTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTC
CTGCTGGACACGAGCTTCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCTGCACTTCTCCTGCTCACCTGCCTTCTTCTGGATGGGCTCGAGGGGTACAACCTTACCAGCTCGTG
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCTCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAATATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCTTTCATGTGCTGGATCCGGGACTCCCTGGTCACTACATCAACCTGGGCCCTCTTCAGC
CTGGTGTCTTCTGTTCAACATGGCCATGTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCCACACCCAA
AAGTGGTCACATGTGCTGACACTGTGGGCCCTCAGCCTGGTCTTGGCCTGCCCTGGGCCCTTGATCTTCTTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTGCTCCTCTACCTTTTCAGCATCATCACTCCTTCCAAGGCTTCTCTATC
TTTCTCTGTTACTGTGCTCATGCGGCTGCAGGCCCGGGGTGGCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATCTAGGCCTCCAGCCCACTGCCCATGTGATGAAG
CAGAGATGCGGCCCTCGTGCACACTGCCTGTGGCCCCCAGCCAGGCCCAGCCCCAGGCCAGTCAAGCCGAGACT
TTGGAAGGCCAAACGACCATGGAGAGATGGGCCGTGGCCATGGTGGACGGAATCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGACTACTCGGCTCTCACTCAGCTCCCAGGGACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCCTGGTCTCTCCTTACAACCCCTGGGCCCGGCCCTATTGCTGGG
GGCCAGGCCCTTGGATCTTGAAGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCAACCTGAGGGCACTCTGCATCCTCTGTCACTTTAACCTCAGGTGGCACCCAGGG
CGAATGGGGCCAGGCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCCTACCTCTGAGCCAGGCCCTCCTCCTCCTCAGCCCCCAGTCTCCTCCTCATCTTCCCTGGGGTTT
TCCTCCTCTCCAGGGCTCCTTGTCTCCTTGTTCAGAGCTGGGGTCCCCGATTCCAATGCTGTTTTTTGGGGA
GTGGTTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTTGTCTACTGCACAAGCCTCGGCCTGCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCTTTGTATGAGCTGCATTGCCCTTG
CTCACCCTGACCAAGCACAGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAATGTGGA
CCATGCCAGTCCCGTCTGGTTTTCATCCCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCCTGACACTCTCCTAAGAGGTTCTCTCAAGCCCCAAATAGCTCAGGCGCCCTCGGCCGCCATCATGGT
TAATCTGTCCAACAAACACACAGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACTCCTCTGCCAACATTAGTCTGGTATGTGAGGCGTGCCTGAAGCAAGAATCCTGGAGCTACAGGGACA
GGGAGCCATCATCTCTGCTGGGAATCCTGGAAGACTTCTGCAAGAGTCAAGCTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTACGTACCAATTCTTTTGTCTTTTGTATATAAAAAGAAGTACATGTTTATTGTAGAGA
ATTTGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAAA
AA

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACCTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT
GGTGTCTTCTGTTCAACATGG

09978299-107504

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTGCTTTGA
 TCCTTTTCAAAAACTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAACTACCCCT
 GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCAGTCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
 TCGGGAGTCCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCACCCAGTCAGCCAAATGAGCCTCTTCGGGC
 TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
 AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
 GAAGTATTACAGCCCCAAGGTTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
 AGGAAAATGTATGGATACAACTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
 ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACATATATAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
 GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTTCTTCTGAACCAGGGT
 TCTGCATCCACTACAACATTGTCTATGCCACAATTCACAGAAAGCTGTGAGTCCTTCAGTGCTACCCCTTCAGCTT
 TGCCACTGGACCTGCTTAATAATGCTATAACTGCCCTTAGTACCTTGGAAGACCTTATTCGATATCTTGAACCAG
 AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTTGTTTTTGAA
 GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
 CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
 GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
 ACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACAAATCACTACCCGACGTGGCCCTGGAGC
 ACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGCGCATCACCAAGCAGCTCTTGCCCA
 GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGCTTATCTCCATCCTTAATCTCAGTTGTTTGCT
 TCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
 ACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAGAAAATTAATGTTGTAT
 TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCTAGCTGGGTTCTGTATTTTCAGTTCTTTC
 GATACGGCTTAGGGTAATGTCAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCCTTGCTTAAC
 TCTAAAGCTCCATGTCTCTGGGCCTAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTTTTGCTCATATTCACAT
 ATGTAAACCAGAACATTCTATGTACTACAACTGGTTTTTAAAAAGGAACATATGTTGCTATGAATTAACCTGTG
 GTCATGCTGATAGGACAGACTGGATTTTTTCATATTTCTTATTAATAATTTCTGCCATTTAGAAGAAGAGAACTACA
 TTCATGGTTTGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
 TTTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTAAATATATCT
 ATTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCT
 AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATGTTGCTCTGACAAAAATACATGTATTTCA
 TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAAGTGAATTGGAATAGAATTGGTAAGTTGCAAA
 GACTTTTTGAAAATAAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
 AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
 AAAGCACCTTGAAAAGACTTGGCAGCTTCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA
 TTGTGATGTTGTGGTTTTATTATCTTAACTCTGTTCCATACACTTGTATAAATAACATGGATATTTTTTATGTACA
 GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAATATTTTTGCTTGT
 AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAAATAAAGA
 ATGTGGCTATTTTGGGGAGAAAATTAATAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTGCTTTGA

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
AGACTAAAAATGGTGTTCCTAATGTGGACACTGAAGAGACAAATCTTATCCTTTTAAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCTTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTTCCACGAACACCACGAACCTCACCCCTC
ACCATTAACCACATACCAGACATCTCCCCAGCGTCCTTTTACAGACTGGACCATCTGGTAGAGATCGATTTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAAAATCCCTTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTATGTTTCATATTCAATAGAGAAA
GATGCCTTCTTAACTTGGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATCTTGACCTAAGTGGAAATTGCCCTCGTTGTTATAATGCCCATTTTCTGTGCCCG
TGTA AAAATAAATTTCTCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAGTTTTACGTCTA
CACAGTAACCTCTCTCAGCATGTGCCCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACCTGGATCTGTCC
CAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAACCTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGTCTTTAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAATTGCTAACCTCAGCATGTTTAAACAAATTTAAAAGA
CTGAAAGTCATAGATCTTTTCACTGAATAAAAATATCACTTCCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCTTGAACAATTACATTAATTCAGATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTTATGTCTGTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCTCTGATTTTCAAGCATCTTTCTTTCTCAAATGCCTG
AATCTGTCTAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTCACCTTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTTGAAGAGCTTCACAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTAATTTCAATCAGAAGGAATTACTCATATGCTAACTTTACCAAGAACCTAAAGGTT
CTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTAATCAAG
AATCTGCTAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTCTTGCCTTCTGGAGTTTTTGTATGGT
ATGCCCTCAAATCTAAGAAATCTCTCTTTGGCCAAAATGGGCTCAAATCTTTTCAAGTTGGAAGAACTCCAGTGT
CTAAAGAACCTGGAACCTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAAGTGTTC
AGAAGCCTCAAGAATCTGATTTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAAGAAATGTCTC
AACAACTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTTCTCATGGTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAGAGAGAAA
CATTTTAAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAACCTTTCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAGACTGAAAATTTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAAG
TCCAAGTTCTCCAGCTCCGAAAAGGCTCTGTGGGAGTTCTGTCTTGTGAGTGGCCAACAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAGAAGCGCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTCTTTGCAAAAACAACTGCCTAGTTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYR
NPCYVSYISIEKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKLMMDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNLSLFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDVWVFWVWNHTEVTIP
YLATDVTVCVGPAGAHKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLLE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSPCDEKKQNDSVIAECSNRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLDDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTLQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDLEFNLYLVGEIVSGAFLTMLPRLEILDLSFNLIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLSLNLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDL
SNLLKTINKSALETKT'TTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYALQRLMDENMDVIIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLEFWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGTGGCAGTGGGCGGCACAGAGCAGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCTCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCTTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTCTG
GAGGAGCAGCTGGGGTCTCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTGA
AAAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

09978259-101501

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTTTCAGCTTGGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGCCCCGGGGAGGGGAACTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTC
TCCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCCTGTCTCCGAGTCGTTCCG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGACGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCTTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTTCTGGAGGAGCAGCTGGGGTCTTGCTCCTGCAAGAAAGA
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCCTCCTCCCCCTTCCTCGGGAGGCTCCCCAGACCCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTL CVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

09473299-101301

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCCTTCCTGGAGGAGCAGCT
GGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTCTTCCTCCTCCCCCTTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAATICPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

099769-10121

FIGURE 218

GGTTGCCACAGCTGGTTTATAGGGCCCCGACCACTGGGGCCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCTGCCGTGAGCTGCCGGCCG
AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCTGGTCTCCCTTCTCCGCTGGGCCGGTTTATCGGGAGG
AGATTGCTCTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAACTCCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCGCGTCATGATGGCCCGGCAAAAGGGCATTCTTCTACCTGACCCTTTTCTCATCCTGGGGACATGTACAC
TCTTCTTCGCCCTTTGAGTGCCGCTACCTGGCTGTTTCTGCTGCTCCTGCCATCCCTGTATTGTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCTGAGTACCCTGGAGTGATTCTCGGGCGCTACCAGATG
AAGCAGCTTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA
TCAAGAATTTCCAGATAAAACAACCAGATTGTGAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGGCCTCCCATTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTG
TTGGAAAGAGGAACTACCGTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTGGCTTCTTGAGACATTGAAAGAACTCCTGGAACTG
TTCTAGAAGTCCCTCATTTGCTTCTTTTACACTCTGGTCCGTCTGTTGGGACTGACTGGATTTTCATACTTTCTCTGCTG
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA
TTTTGCCACTGGAGGAAAGTGAAGTCAACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCAAGGAGGAGCTGAAGCTGAGAAGT**AG**CCCTATCTATGGAAGAGACTTTTGTGTTGTGTT
TAATTAGGGCTATGAGAGATTTTCAAGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTAGTCACCAGTTGCACACTGGCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACT
CAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCACAGGGTCCCTTGGACCCCCCTCTCTCCCTCCAGATCCAGCCCTCCTGCTTGGGGTCA
TGGTCTCATTCTGGGGCTAAAAGTTTTTGGAGCTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATCTCTGGCCACCAAGTGCCAGCATTGCCACAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTCTTTTACTCCTGCTCCCATTAGGAG
CAGGAATGGCAGTAATAAAAAGTCTGCACTTTGGTCATTTCTTTTCTCAGAGGAAGCCCCGAGTGCTCACTTAAAC
ACTATCCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAAGTTCTTCCATTAAGCCT
CGGCTGAGTGAGGGAAAGCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCACCTCTGGCT
ATGGTAACCACACTGGGGCTTCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTT
CACCTTGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTAGGGAAGAAGATTTATGT
ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT
GCGGTGGGGAGTGTAACCGGAACTTTTTCATCTATTTGAAGCGGATTAAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLTIIYVFAFNIVYVALKSLKIGFLETCLKETPGTVLEVLIFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

0903209 101509
100707 0029060

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCAACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACCTATCCACCATA
TTTCAGAAGCCAAGGTACAGGTGTCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCAGAGTGAGCAACGGCACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGACAGATATTTAGGTACAAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCCAGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMD NVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNV DVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYEC SASNDVAAPV VRRVKVTVNYPPYI SEAKGTGVPVGQKGT LQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLI FFNVSEHDYGN YTCVASNKLGH TNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

05973290 "101501
TGGT" 6682680

FIGURE 223

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

099359-101501
"0570" 652660

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCGACTCCACC**ATGA**ACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCAAGCACCTGCCTTACA
GAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCTGTGAGGACTTTTAC
CAGTTCTCCTGTGGGGGCTGGATTTCGGAGGAACCCCTGCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCTTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTTACCCTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCAGTCTGGGCTCTTCTGCCCCTCTCGGGATTACTACTTAAACAGAAGTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAAACATCACAGTGCCCCAGGACGAGCGGCGGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCAACATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGTCAAGCTCATCAACCGCACGGAACCAAGCATCTGAAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTCTTCCAAAACATGTTGAATTTGTACAACCTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACTACCTTCCAACTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTCGGTGGCATC
GGTGTGGTCAATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAAACATTACTGACAAACGGGGGGCTGAAG
GCTGCCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCTGGGATTGCCCAGGTGTGGTGTCTGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGTGGGCACTCTCTCCAACCTCCCGTGACTTCTCTGCGG
CACTTCGGCTGCCCTGTGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGG**TAG**ACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCACAGGTGACATGAGTACAGACCTCCTCAATCACCACATTG
TGCCTCTGCTTTGGGGGTGCCCTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTACCCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCTATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFT VYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANЕКVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTA FEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGLCEVW

Type II Transmembrane domain:

amino acids 32-57

0597829-10101
T03T0T" 66287560

[illegible]

GCCGCGCCCTCCGCCCTCCGCACTCCGCGCTCCCTCCCTCCGCCCGCTCCCGGCCCTCCTCCCTCCCTCCCTCC
 CAGCTGTCCCGTTTCGCGTCATGCCGAGCCTCCCGGCCCGGCCCGCCGCTGCTGCTCCTCGGGCTGCTGCTGCT
 CGGCTCCCGGCCCGGCCCGCGGCGCCGGCCAGAGCCCCCGTGTGCCATCCGTTTCTGAGAAGGAGCCGCTGCC
 CGTTCCGGGGAGCGGCAGGTAGGTGGGCGCCCGGGGGAGGCGCGGGCGGGGAGTCGGGCTCGGGCGAGTCAGCGC
 CAGCCCGGAGGGGGCGGGGCGCAGGTGGCTCGGCGCGCGCGGCCCGGAGGGTGGGCGGGGGCAGAAGGGC
 CGGTCGCTGGGACCCGGGACCCGCGGGCAGCCCCGGGCGGCACACGGCGCAGCTGGGCGAGCGGCCTCAGC
 CAAGCCCGTCCCCGAGGCTGCACCTTCGGCGGGAAGGTCTATGCCCTTGGACGAGACGTGGCACCCGGACCTAGG
 GGAGCCATTCCGGGTGATGCGCTGCGTGCTGTGCGCCTGCGAGGCGCAGTGGGGTCGCCGTACCAGGGGGCCTGG
 CAGGGTCAGCTGCAAGAACATCAAACAGAGTGCCTAACCCCGGCCTGTGGGCAGCCGCGAGCTGCCGGGACA
 CTGCTGCACAGCTGCCCCAGGACTTCGTGGCGCTGCTGCACAGGGCCAGGTCGACGCGGTGGGCACGAGCC
 AGTCTCGTGTCTGCTGCTCTAGCCTCCGCTTCTCTATCTCTACAGGCGGTGGACCGCCCTACCAGGATCCGCTT
 CTCAGACTCCAATGGCAGTGTCTGTGTTTGGACCCCTGCAGCCCCACCCAAAGATGGCCTGGTCTGTGGGGTGTG
 GCGGGCAGTGCTCGGTTGTCTCTGCGGCTCCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCA
 CCCTTCAGGGGAGGTCTGGGGGCCCTCTCATCCGGCACCGGGGCCCTGTCCCAGAGACCTTCAGTGCCATCCTGAC
 TCTAGAAGGCCCCACAGCAGGGCGTAGGGGGCATCACCTGTCTCACTCTCAGTGACACAGAGGACTCTTTGCA
 TTTTTGTGCTCTTCCGAGGCTTGCAGGACTAACCCAGGTTCCCTTGAGGCTCCAGATTCTACACAGGGGCA
 GCTACTCGAGAACTTCAGGC CAATGTCTCAGCCAGGAAACAGGCTTTGCTGAGGTGCTGCCAACCTGACAGT
 CCAGGAGATGGACTGGCTGGTGTGCTGGGGGAGCTGCAGATGGCCCTGGAGTGGGCAGGCAGGCCAGGGCTGCGCAT
 CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGTCTTGCAAAGTGTCTTTGTGGGGCTAATGCCCTGATCCC
 AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGTCTAGGAAATGGCNCCTGATCCTCCAGGTGCAATT
 GGTAGGACACCAAGTGAAGTGGTGCCATGACACTGGAAACCAAGCTCAGCGGAGGGATCAGCCCACTGCTCT
 GTGCCACATGGCTGGCCTATCTCCTCGCCCCAGGCCCTGGTATCTGCCCTGGGCTGGGGTGCCGAGGGG
 TCATATGCTGCTGCAGAAAGAGCTCTTCTGAACTGGGGCACCAAGGACTTCCAGACCGAGAGCTTCGGGGGCA
 ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCCCTGCCCGTGCCCTTAGCAGGAGCCCTGGTGTACC
 CCCTGTGAAGAGCCAAAGCAGCAGGGCAGCCTGGCTTCTCTTGGAATACCCACTGTCACTTGCATGAAGTGTCT
 CTGGTCTGGGCTTGGTGCTCAGAACAGGCACCTGTCACTGCCACCTCCTTGGGCCCTCTGGAACCGCAGGGCC
 TCGCGCGCTGCTGAAGGGATTCTATGGCTCAGAGGCCAGGCTGTGGTGAAGGACCTGGAGCCGGAAGCTGCTGCG
 GCACCTGGCAAAGGCATGGCTTCCCTGATGATCACCCACCAAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCC
 CTCTCCAGGTGCACATAGCCAAACCAATGTGAGGTTGGCGGACTGCGCCTGGAGGCGGCCGGGGCCGAGGGGGT
 GCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCGCCAGCCTGTGTGCTGCCCTGCCCTCCCGCCCTAGCGCCGC
 CAACCTGGTGGTCTCTGGGCGGCCCGAGACCCCAACCACTATGCTTCTCGAGGGGACAGCGCCCCCAGGGGG
 CTCTGGCGGCCCAACTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACCAAGGTGATCTGTGACCCGGT
 GGTGTGCCACCCGCCAGCTGCCACACCCGGTGCAGGCTCCCGACCAAGTGTGCTGCCCTGTTTGGCCTGGCTGCTA
 TTTTGATGTTGACCGGAGCTGGCGGGCAGCGGGTACGCGGTGGCACCCCGTTGTGCCCCCTTTGGCTTAATTAA
 GTGTGCTGTCTGCACCTGCAGCAGGGGGGCACCTGGAGAGGTGCACCTGTGAGAAGGTGCAGTGTCCC CGCTGGC
 CTGTGCCACGCTGTGCTGTCAACCCCAACGCACTGCTGTCAACAGTGTCCAGGTGAGGCCCAACCCAGCTGGG
 GGACCCCATGCAAGCTGATGGGCCCCGGGGCTGCCGTTTTGCTGGCGAGTGTGTCCCAAGAGTCAAGAGCTGGCA
 CCCCTCAGTGCCCCCGTTTGGAGAGATGAGCTGTATCACTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGT
 GTGAGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCAGGGAGGTCCCTGAAGAAGCTGAAGGTCA
 CTGTGTCCAAGTGCCTCTGGGGGACACTCAGTGTCTGCTGTCTTGTACCAGGCAAGGGGTGCCTCACTGTGAGC
 GGGATGACTGTTCACTGCCACTGTCTGTGGCTCGGGGAAGGAGATGCTGCTGTCTCCGCTGCACGGCCCCAC
 GCGCGCTTAAGTGGAGGAGTCCAGGCTCAGCAGCTGTGAGTGGAGGGCTCACTGCTGTGGGACTCCTGATCAG
 GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACAAGTGCCATTCCAATCCACCTCACAGCA
 ACCTGGTGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTTCTGAAGCTCAGAGAAATTAAGCAACGAG
 ATGAAGTACCCAGCTGTGTGCACTGACCTGTTTAGAAAACTATGGCCTTTCTGGGACCAAGGCAGGGATGCTT
 TGCCCTGCCCTCTATGCTCTCTGTGCTCTCCACTCCCTCCTCCCACTCCTCCCTCCTCTGTCTCTC
 AGAGCCCAAGAGACGAGCAACTGATCAGAGCTGGAGAAAGAAAGCCGAAGCTTCTTAGGGAGCAGCCAGAGGGCC
 AAGTGACCAAGAGATGGGGCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCAATTCTCCTGTGGGAAG
 CCCAGTGCCTTTGCTCCTCTGTCTGCTCTACTCCACCCCCACTACCTCTGGGAACCAAGCTCCACAAGGGG
 GAGAGGCAGCTGGGGCCAGACCGAGGTACAGCCACTCCAAGTCTGCCCTGCCACCTCGGCCTCTGTCTGGAA
 GCCCCACCCCTTCTCTCTGTACATAATGTCACTGGCTTGTGGGATTTTAAATTTATCTTCACTCAGCACCAAG
 GTCCTCCGACACTCCACTCCTGTGCCCTGAGCTGAGCAGATCATTAATTGGAGAGTTTGTATTATTAAAC
 ATTTCTTTTTCAGTCTTTTGGGCTAGAGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCTGGTGGAGAAGGG
 GCNAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG
 CGTGGCANNITGGCTGGCATNCCTGGGTTCCGCAGAGGGGGCTGGGGATGGTCTTTGAGATGGTCTAGAGACTCAAG
 AATTTAGGGGAAGTAGAAGCAGGATTTTGACTCAAGTTTAGTTTCCACATCGCTGGCCTGTTTGTGACTTTCATG
 TTTGAAGTTGCTCCAGAGAGAGAATCAAAGTTGTACCAGCCCCCTCTCTCCTCCTTCCCTTCCCTTCCCTTTCT
 TTCCCTCCCCTCCCCTCCCCTCCCCTCCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCACAGCACCTGCCTTACAGAGGCCTGCATTTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTTCGGA
GGAAACCCCTGCCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCCCT
CCTTTCTTTCTTCTTTTCTTCCCTCCCTCCCTTTTCTTCCCCTTTTCTTCCCTTCTTCCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTTCG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTTCAGTGTGATGGGTTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAAC TAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCC
TTCTATGCCCCGAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCAACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCAATTGTGCCTCTGCTTTGGGGGTGCCCTT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCTTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

0997229 "101501
T095T0T" 6624660

FIGURE 229

CCCACGCGTCCGAGCCGCCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAACAACACCCAACAACCTGGGGTGG
GGGGAAGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCCGGGAAGACTCGCCAGCACACAGGGGG
TGGGGGAGTGCAGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGGTGCAGGGTGCTTGTGCTCGAACCACTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCTGGCTGAACCGGTCAAGTATTATTTTGCGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTTCACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACAGGCTGGGGAATATGAATGCAGTGCAGGAAATGCTGTGTCATT
CCCAGATGTGAGGAAAGTAAAGTTGTTGTCAACTTTGCTCCTACTATTTCAGGAAATTAAT
CTGGCACCCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTCACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTTAAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTTCATGTTTGTAACTTACAACCTTTTCAAAGCATTTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTTCAGTTCCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG
CATTTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTTCAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGTAAGTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTGTCTCCAATAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATAACCATATGATTTCT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCTTTTAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTTCACTAGTCCAAGCCAAAA
TTTTAAGATGATTTGTGAGAAAGGGCACAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACCTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGD₁TAVLR₂CYLE₃D
GASKGAWLNRSSIIIFAGGDKWSVDPRVSI₄STLNKRDYSLQIQNV₅DVTDDGPYTCSVQTQHT₆P
RTMQVHLTVQVPPKIYDISNDMTVNEGTVN₇TLTCLATGKPEPSISWRHISPSAKPFENGQYL₈
DIYGITRDQAGEYECSAENAVSFPDVRKVKVVVNFAP₉TIQEIKSGTVTPGRSGLIRCEGAGV₁₀
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQE₁₁HFGNYTCVAANKLGT₁₂TNASLPL₁₃
NPPSTAQYGITGSADVLFSCWYLVLTLS₁₄SFTSIFYLKNAILQ₁₅

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTTCACTCAGAACCAGGGATGTTTCAAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGGACCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCCAAAAGGGGAACTGGAAC TAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTGCGGCTCCTGGAGCCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCCACCAGGTTTGGCTGGTTTTCTTGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGCCCTCCTGGTGAGTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC
GTCAGGATTGTGCGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCAGTGTGCGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCGGAAACCCCTTTCA
CTTCTCTGCTCCCGAGGTGTCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

66234660

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCCCAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCCAAAGACGTTTTCTTTGGACCAAAGATCTCTTTTCG
TGATTCCTTGCAACAATCAATGAGAATCTTCATGTATTCTGGAGAACACCATTCCCTGATTC
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

105707-653600

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG

KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

10978299 101501
105197 652460

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
 CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC
 GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
 TGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCCGCGCTGGCTGTG
 CGCTGGGGCGCTGGTGTGGCGGGTGGCTTCTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA
 TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTGGAT
 GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
 AGGAACAGAAACAAAACCTTTAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTGGCC
 TGGATTCTGTTGAGCTAGCTCATTATGATGTCTGTGTCTACCCAAATAAGACTCATCCC
 AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
 ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTCTCTCCTC
 AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAA
 TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
 TTTAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
 CCGACCCTGCTGACTCTTTGCTCCTCTGGGTGAAGTCTTATCCAGACGGTTGGAATCTTCTCT
 GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC
 AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGAATTCAGAGGCTGTGGTCTTCCAA
 GTATTCCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCTAGAAAAAATGGGTGGC
 TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
 CTTTACTGGAACTTTTCTACACAAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
 CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
 CTGGGAGGTCAACGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGAGCAGCTGT
 TGTTTCATGAAATTGTGAGGAGCTTTGGAACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
 CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGCA
 GAGGAGAATTCAAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
 AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
 TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
 ACTAAAAAAGTCTTCCCCAGAGTTCAGTGGCATGCCAGGATAAGCAAATTTGGGATCTGG
 AAATGATTTTGGAGGTGTTCTTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTA
 AAAATTGGGAAACAAACAAATTCAGCGCTATCCACTGTATCACAGTGTCTATGAAACATAT
 GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG
 AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTTATTGTCGAGATTATG
 CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
 ATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
 TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
 TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTTATTGATCCATTAGGGTTACCAGAC
 AGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
 ATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCCT
 GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
 TTGAGTGAAGTAGCTTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
 CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAAAATGGTATATTTGAAATAAAGT
 TGAATATTATATATAA

09978299 " 401504
 10978299 " 401504

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIRYGVFRGNKVKNQAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQQRGNILNLNGAGDPLTPGYPANNEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHVSVEYET
YELVEKFYDPMFKYHLLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDLSFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713